Water quality assessment in coastal lagoons using improved back propagation deep neural network with PH level compensation sensors in smart environment

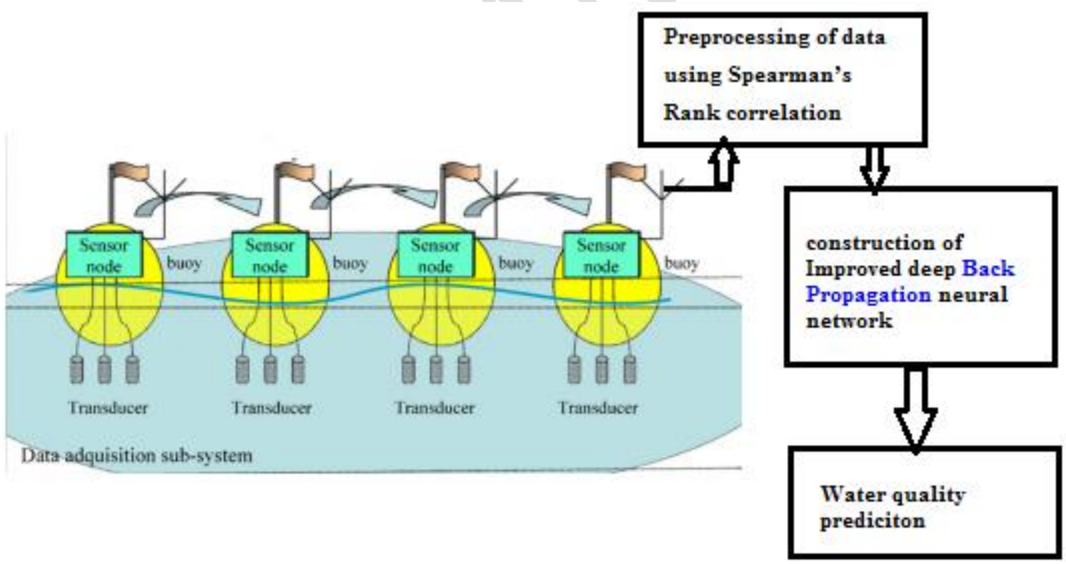
VENKATESAN. V[1*], ADILAKSHMI. A[2]

Department of Civil Engineering, University College of Engineering, Anna University, Ariyalur-621 731, TAMILNADU-INDIA, venkivenkatesan367@gmail.com/1vvn6030@gmail.com

Department of Sciences & Humanities, University College of Engineering, Anna University, Ariyalur-621 731, TAMILNADU-INDIA, adiucea@gmail.com

Corresponding Author: VENKATESAN. V[1*]; 1vvn6030@gmail.com

Graphical Abstract:



Abstract:

In coastal lagoons, eutrophication is a significant ecological and environmental problem. The majority of the pollution and deterioration problems that coastal zones and their ecosystems face are due to human influence through urbanization and industrialization. In this research, evaluations of water quality were made throughout a few lagoons in India during both the wet and dry seasons to establish their suitability for harboring aquatic life and the impact of seasonality on their quality. In order to reduce the effect of environmental influences on sensor based on nanoparticle sensitive film, an enhanced deep Back Propagation (BP) neural system was developed. The incompatibility between high performance and low calculation time was typically found in today's modern sensor compensation techniques. Moreover, higher recognition errors were always the outcome of weak self-learning abilities. To solve these problems, a 16-layer deep BP neural network model was developed after hyper parameter searching. In Addition, the stochastic gradient descent (SGD) method and the mini-batch method were employed to achieve a proper balance between model performance and training time difficulties. The enhanced deep BP neural network was a great compensation solution for the sensor used in complicated environments due to its self-learning and self-adaptability. The proposed method is named as Imp_BPNN_SGD is analyzed in terms of various parameters and found that it achieves 35.6% of RMSE, 21.6% of MAE, 19.6% of RAE and 97.2% of D^2

Keywords- water quality, smart environment, neural network, coastal lagoons, Eutrophication, backpropagation.

Introduction:

Ecotones called lagoons form where coastal, terrestrial, and marine environments converge. Lagoon systems make for 13% of the planet's shoreline globally [1]. Fresh and salt water combining creates coastal lagoons [2]. They are dynamic, heterogeneous systems that incorporate fresh water as discharges from streams or other waterways and also sporadic or continual waves of seawater which slip through sandy barriers [3]. The groundwater affects the water's ionic concentration in certain lagoons. Each of these coastal regions is a special, delicate environment that is very important as a natural habitat for many different living types. In terms of prospects for agriculture, aquaculture, tourism, and leisure activities, lagoons are also crucial to the socioeconomic framework [4]. However, if the aforementioned socioeconomic benefits are not carefully regulated and supervised, they could become harmful to the ecological environment. Lagoons are vulnerable to anthropogenic pollution and have restricted water circulation to account for fluctuations in water quality [5]. In coastal lagoons, monitoring water quality is crucial for maintaining ecological characteristics. Freshwater body water quality is a complicated topic that involves many factors, including physically, chemically, and biological processes [6]. Recently, a variety of tools, such as numerical analysis, optimization techniques, and hybrid decision support methods, are invented to assist with water quality management [7]. Quality of water indices are developed and utilized globally in addition to the expanding use of these advanced tools due to their simplicity, versatility, and user-friendliness. In order to monitor the water quality in coastal lagoons experiencing environmental issues such eutrophication, organic pollution, and growing stress on aquatic biota, a back propagation deep neural network method is developed.

The research contributions are as follows:

- To analyze the hourly high efficiency variation and provide long-term simulations of pore water levels, an enhanced deep back propagation neural network is developed.
- Results from stochastic gradient optimization are achieved by using only experimental data for the affecting external factors, rainfall and evapotranspiration as inputs and the past forecasted values of pore water heads as recursive inputs.

The organization of this paper is: water quality monitoring using various neural network is discussed in section 2. The suggested feature extractor and predictor with optimizer are elaborated in section 3. The performance analysis of proposed method is examined with comparison of existing methods in section 4. Section 5 concludes the overall analysis of proposed method.

Literature Survey:

In [8] the author provides a thorough examination of recurrent cell and network models in order to enhance eutrophication procedures in the Venice lagoon. Additionally, the trained models' potential for transferability was examined. Based on empirical analysis, recurrent neural network models reduce the computational difficulties and also have better accuracy for predicting eutrophication. In [9] the author presented a water quality analyzing technique based on a novel geographically neural network weighted regression. This method achieves high accurate and water quality depends on the significant index of Water Quality Classification Standards Benchmark

In [10] the article objective is to forecast the dissolved oxygen (DO) in the water dumped into this lagoon. Additonally, a multiple linear regression (MLR) based back-propagation neural network (BPNN) methods are used for theoretical analysis. Finally, the accuracy of the predictions provided by the neural network and the other theoretical models were evaluated using the parameters of coefficient of determination (R2), the root mean square errors (RMSE), and the mean absolute error (MAE). In [11] the author developed a feed forward neural networks (FFNs) to provide extend period simulations of hourly pore water levels, in a coastal unrestrained aquifer placed in the Lagoon of Venice. The results reveal that the constructed FNN can accurately recreate shallow aquifer pore water depths over several months. In [12] to calculate the total organic matter (TOM) content of sediments, the author investigated a number of statistical techniques, including cluster analysis (CA), principal component analysis (PCA), partial least squares (PLS), principal component regression (PCR), and ordinary least squares regression (OLS). The TOM can calculated by the parameters like temperature, dissolved oxygen (DO), nitrite (NO2), pH value, electrical conductivity (EC). The most important factors in the lagoon's water quality variability, according to the findings, were the nutrient concentration and the DO gradient.

In [13] the author developed a unique cost efficiency method depends on the geo-ecological information-modeling system (GIMS) to analyze the problems in water quality monitoring of a coastal lagoon. Furthermore, predicted values from the simulation experiments and the real value gathered from field observations are combined to perform this proposed method. This makes it possible to optimize the monitoring system and record lagoon water quality with reliability. Elman neural network (ENN) model was utilised by Liu et al. [14] to forecast DO for quick assessment of Singapore coastal waters. Seven to eight hidden layer nodes made up the network architecture that was planned for this investigation, and positive DO outcomes were attained.

In [15] the author developed a multilayer perceptron-back propagation (MLP-BP) based artificial neural network (ANN) method to identify the level of eutrophication presence in the water. The parameter used for this experiments from Florida Bay water quality monitoring stations (FLAB03 and FLAB14). The monitoring and forecasting of algal blooms can be done using the MLP-BP technique, which is important for coastal lagoons. In [16] WQIs employ a statistical evaluation based on pre-established threshold values established by organizational bodies. A parameter set, relative weights, normalization slopes, and aggregation techniques serve as the foundation for subjective WQIs. Machine learning techniques were used by Jimeno-Saez et al. [17] to calculate the chlorophyll in saltwater from the Mar Menor coastal lagoon. Support Vector Regressions (SVRs) were the techniques employed; the SVRs produced good validating results for the forecasting of chlorophyll concentration.

Water quantity (such as flow and rainfall-runoff) prediction has received greater attention in prior reviews about ANNs than water quality prediction. (for instance, suspended solids [SS]), and river systems were the main scenarios they looked at. In addition, prior assessments have focused on the model's evolution while ignoring the output strategies between the inputs and outputs in a certain prediction tactic. In order to overcome the limitations above, this research focuses on the use of new neural network methods for water quality prediction, with more water quality variables

System model

The depth node, current metre node, water stream node, and full node are the four main types of sensor nodes considered by the network. The water stream node is a GPRS node, whereas the other nodes in the WSN are ZEDs. Depth nodes collect data on water level and temperature through pressure sensors via temperature sensors situated at the sea's bottom, rspectively. Current speed and direction are measured by current metre nodes. Full nodes use temperature and salinity probes through the buoy's line to evaluate temperature, salinity, sea level, and current profiles. In addition to the fundamental physical parameters are temperature, pressure, and salinity, the waterstream node also monitors turbidity, dissolved oxygen, chlorophyll, and nitrates. The final node concludes the WSN's hydrodynamic analysis of the three inputs, as displayed in figure 1.

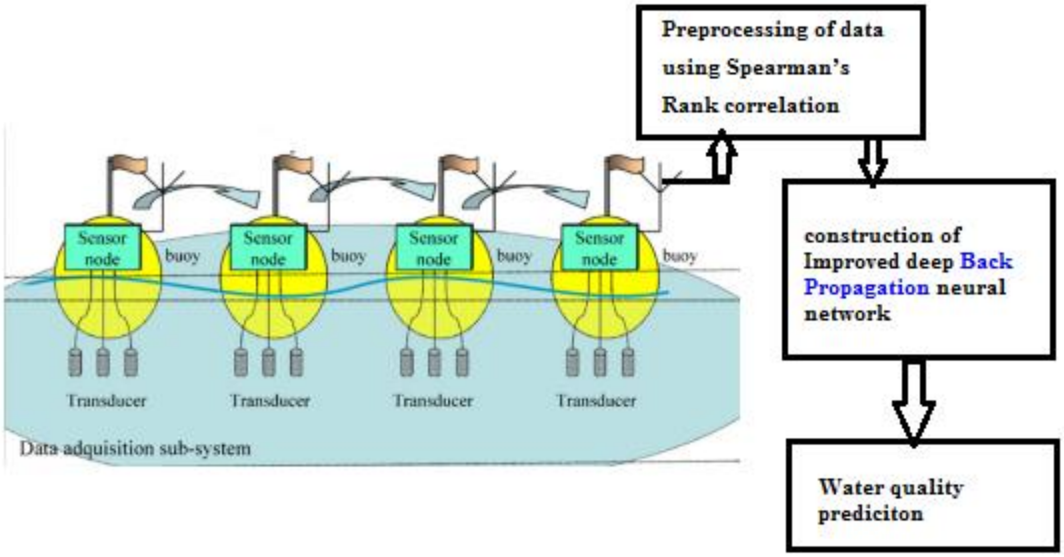


Figure-1 data collection from coastal lagoons using sensor nodes

Case study

On the eastern coast of India, Pulicat is the second-largest transitional water environment after Chilika. Its soils are coastal alluvium and it is part of the hot sub-humid to semi-arid ecozone. The lagoon is typically between one and two metres deep and covers an area of about 450 km2. It's been reported that the lagoons has gotten shallower over time. The system's narrow (200–500 m) southern entrance, which connects it to the Bay of Bengal on its eastern side, is where water primarily moves between the lagoon and the sea, with the exception of occasions when it is completely closed. The exchange through a smaller entry (between 50 and 150 metres) towards the north, which lacks a direct link to the lagoon, is very constrained. As a result, we do not consider this characteristic to be a substantial influencer of salinity dynamics. Two rivers—the Araniar and the Kalangi, which are situated in its southern and northern halves, respectively—bring in freshwater along its western bank during the monsoon season (October to December), with little or very little flow throughout the dry seasons. At the point where the Kalangi and lagoon confluence, the annual average flow of the stream is about 92 Million Cubic Meters (MCM), whereas the latter is about 100 MCM. The National Jal Jeevan Mission and the Indian Council of Medical Research collaborated to perform water analyses. Electrometric probes were used to measure temperature, pH, and electrical conductivity on-site using internal techniques based on the appropriate Standard Methods such as SM2550-B, SM 4500 H, SM

2510-B and SM 4500-O-G,correspondingly. In contrast, ex situ parameters like nitrates, chlorides, sulphates, and phosphates were measured using an internal process based on UNE EN-ISO 14911 and Ion Chromatography.

Preprocessing of data using Spearman’s Rank correlation

Spearman's rho is equivalent to Pearson's correlation in that it can be used to assess how well two variables are related. This rank correlation measurement is nonparametric. The ability of Spearman correlation to analyse monotonic relationships distinguishes this from Pearson correlation. Assume we have the x_i and y_i series. By using the following equation, one can determine the Spearman's rank correlation coefficient:

$$\boldsymbol{\tau_s = 1 - \frac{6\sum d_i^2}{n(n-1)}} \qquad \textbf{(1)}$$

where x_i is the difference between ranks for each x_i , y_i is data pair and also the quantity of data pairs. A perfect Spearman correlation takes place when there aren't any repeated data and each variable is a pure monotone function of the others. Spearman correlation will approach +1 if the data have a comparable rank. On the other hand, data with a different rank will decrease to a value of 1. Additionally, it will approach 0 if the two series are unrelated.

Improved deep Back Propagation neural network (IDBP)

A subset of neural networks that learn using back propagation is known as BP neural networks. Input, implicit, and output layers make up its three levels of structure. Figure 2 illustrates how BPNN gradually modifies the input weights between layers using the training error. Additionally, a Stochastic gradient descent algorithm is used in this adjustment process.

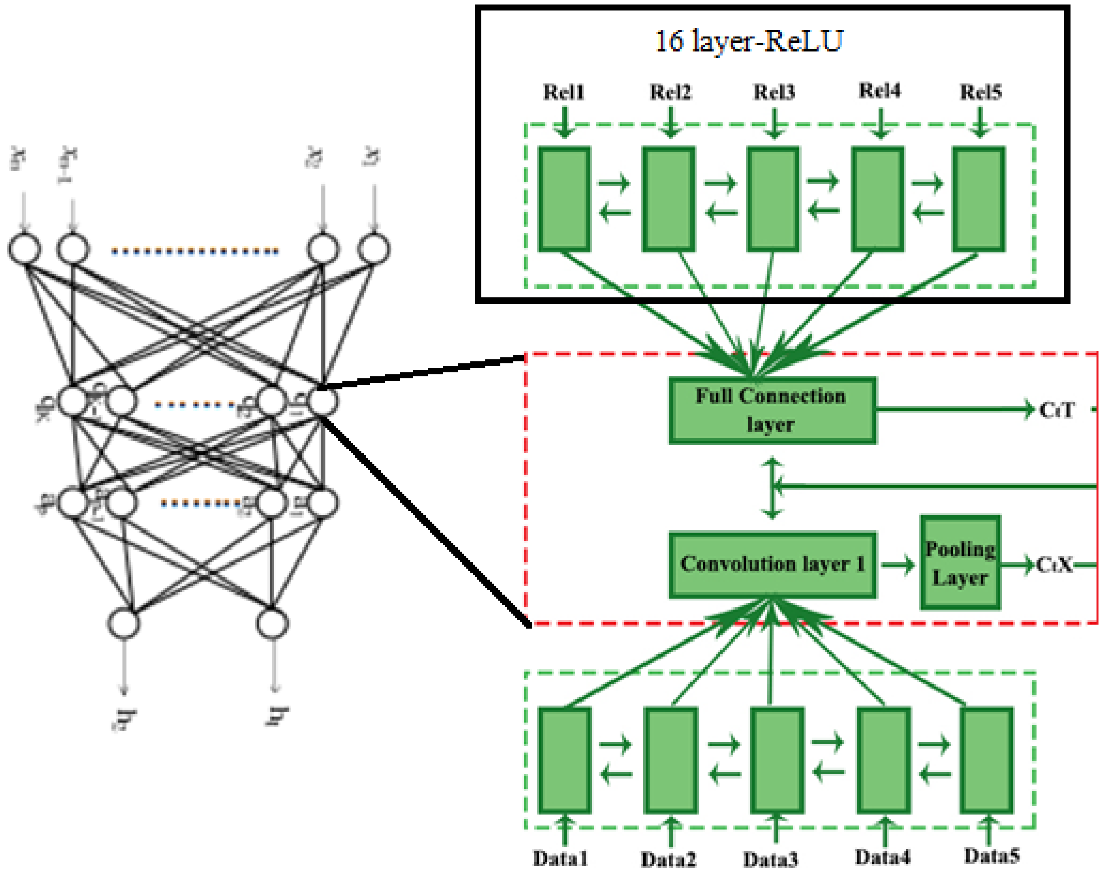


Figure-2 Architecture of IDBP with 16-layers

The network will continuously compute through backpropagation if the output error of the sample does not meet the predetermined convergence error. The connected layers' parameter values are altered by the iterative process, which lowers the error. When the error is brought down to the desired value, the iteration is complete.

The exact computational output equation for this model is:

$$y_j' = S\left[\sum_{I=1}^{P} b_i\, \omega_{ij} + \gamma_j\right] \qquad (2)$$

where b_i is the activation values for the i-th input to the j-th output; ω_{ij} is Connecting weighting coefficients initially involves a small, random amount.; γ_j is the output layer unit threshold; S is the sigmoid function that is $S(s) = \frac{1}{1+e^{-s}}$. The input to a layer m of a IDBP is a $n_1^{(m-1)} \times n_2^{(m-1)} \times n_3^{(m-1)}$ object with $n_c^{(m-1)}$ so $I^{(m-1)} \epsilon \mathbb{R}^{n1(m-1)\times n_2^{(m-1)\times n_3^{(m-1)}}}$ and its elements are denoted by I (m') i,j,k where i, j, and k index the volume and ` selects the channel. The output of a layer m is determined by its dimensions, i.e., $n_1^{(m)} \times n_2^{(m)} \times n_3^{(m)}$ and also the number of filters or channels it generates $n_c^{(m)}$. Convolution of the input with a filter yields layer m's output, which is calculated as $I_{i,j,k}^{(m,l)} = f_{tanh}\left(b^{(m,l)}\right) + \sum_{i,j,k,l} I_{i,j,k}^{(m-1,l)} W_{i-i,j-j,k-k,l}^{(m-l)}$where, $W^{(m,l)}$ and $b^{(m,l)}$ are the parameters where the lth filter in layer m is defined. The size of the filters (i.e., the values of $W^{(m,l)}$) that are non-zero and the areas where the filters are evaluated (i.e., the values of i, j, k for which $I_{i,j,k}^{(m,l)}$ is determined) are factors of the network model. Hence, we employ a hyperbolic tangent activation function. with $f_{\tanh(a)} = \tanh(a)$. ReLu layers maintain the inputs' spatial structure and, as even more number of layers are added, provide progressively more complicated conceptions of the input. A fully connected network layer is then created using the output of the previous layers as its input. This is accomplished by treating the layer's output as a single vector while ignoring the spatial and channel structure. A vector $I^{(m)}$ with a dimension that is a network architectural parameter is the result of a fully connected system. In m layer the output neuron I is provided as

$$I_i^{(m)} = f_{ReLu}(b^{(m,i)} + \sum_j I_j^{(m-1)} W_j^{(m,i)}) \tag{3}$$

where, $W^{(m,i)}$ and$b^{(m,i)}$are the parameters of neuron I in layer m and the summation over j represents an accumulation of all input dimensions. The activation function $f_{ReLu}(.)$ is selected as a Rectified Linear Unit (ReLU) with $f_{ReLu}(a) = \max(0, a)$. The sparsity it causes in the outputs is thought to be mainly useful in classification tasks since it assists in the segregation of classes during training. The output layer receives its input from the last fully connected layer. The specific task determines the output layer's format and structure. Here, two distinct output function types are considered. The softmax function is a typical output function in K-class classification problems:

$$f_i = \frac{\exp(I_i^{(o)})}{\sum_j \exp(I_j^{(o)})} \tag{4}$$

$$I_i^{(o)} = b^{(o,i)} + \sum_{k=1}^{k} W_k^{(o,i)} I_k^{(N)} \tag{5}$$

where, N denotes the final fully connected layer, $b^{(o,i)}$ and $W^{(o,i)}$ are ith output unit parameter and $f_i \epsilon [0,1]$ is the class I output, which is in range of that class given the inputs. In addition, we analyse a variant of the logistic output function:

$$f = a + (b - a)(1 + \exp\,(b^{(o)} + \textstyle\sum_j W_j^{(o)}\, I_j^{(N)})^{-1} \qquad (6)$$

This produces a continuous output f with parameters $b^{(o)}\ and\ W^{(o)}$ that must be within the range (a, b). This is referred to as the scaled logistic output function. The BP neural network's satisfaction is greatly impacted by the transfer function of every layer. Usually, experimental analysis is used to calculate these transfer functions. When the input layer and first hidden layer in the BP neural network have a "tansig" transfer function, a "tansig" transfer function exists between the first hidden layer and next hidden layer, and a "purelin" transfer function exists between the next hidden layer and output layer.

$$h_1 = a_1\left[\textstyle\sum_{i=1}^{p} \tan sig\ [\sum_{j=1}^{k} \tan sig\ [\sum_{k=1}^{m} x_k w_{kj} + b_j] w_{j1} + b_1] w_{i1} + b_1]\right] + b_1 \quad (7)$$

$$h_2 = a_2\left[\textstyle\sum_{i=1}^{p} \tan sig\ [\sum_{j=1}^{k} \tan sig\ [\sum_{k=1}^{m} x_k w_{kj} + b_j] w_{j1} + b_1] w_{i1} + b_2]\right] + b_2 \quad (8)$$

Following a review of the sequential data for any findings that did not correspond to the expected time intervals, observations with inconsistent time intervals were eliminated, and missing values were determined by interpolation.

Stochastic Gradient Descent process

In machine learning (ML) applications, stochastic gradient descent (SGD) is an optimization process used to identify model parameters that are consistent with the best predicted-actual output. When integrated with backpropagation, it is a potent method that is frequently applied in neural network training applications. For memory-related uses, it must have a single training record that is processed by the entire network. The processing of a single record also makes it run more efficiently. It can operate more quickly when dealing with larger datasets since it updates the parameters more frequently. The SGD optimization is presented as the following equation:

$$x^* = \arg\min_{x \in X} f(x) \qquad (9)$$

The goal of this function is to minimise the target score that is used as the validation score; x* is a set of hyperparameters that provide the lowest score, and x can take any value from the X field. The issue with optimising hyperparameters is quite expensive to test the objective role to discover the solution. Then, while experimenting with different parameters, we must estimate the validation metrics after training the model on training data and forecasting validation data. The weights wi that linked the input and hidden layers, as well as the biases $b_{i,}$ were to be initialised and modified in each iteration of the stochastic gradient descent-based learning strategy. Because of this, the training procedure was frequently time-consuming, and the training model occasionally failed to achieve global minima. However, IDBP did not tune those factors; instead, it employed minimal norm least-squares. IDBP with fixed wi and bi was similar for predicting the least square solution during the training process in equation 6 $\hat{\beta}$ of $H\beta = T$

$$\left\|H(w_1, \dots, w_M, b_1, \dots, b_M)\hat{\beta} - T\right\| = \; min_\beta \|H(w_1, \dots, w_M, b_1, \dots, b_M), \beta - T\| \quad (10)$$

with the minimum value $\hat{\beta} = H \dagger T$. where, $H \dagger$ was matrix H's generalised Moore-Penrose inverse. This solution's primary distinguishing characteristics include lower training error, a unique solution, and the smallest weight norms.

Performance analysis

To demonstrate the effectiveness of the suggested Imp BPNN SGD model for predicting water quality, the hourly temperature (T) of water, pH, dissolved oxygen (DO), electrical conductivity (EC), and turbidity (NTU) data, which together total 39,025 features, are taken from the water quality sequence data of Pulicat. These features are listed in table 1

Table-1 dataset features with their statistical analysis values

Feature	**Min**	**Max**	**Mean**	**Standard deviation**	**median**
Temperature	20 degree Celsius	98 degree Celsius	7.684	1.578	7.459
pH	0.223	0.457	2.3	1.78	7.98
Dissolved Oxygen	34.67	87.94	7.459	0.787	6.548
Electrical Conductance	51.6	68.45	6.35	1.578	7.347
Turbidity	1.5	2.89	1.6798	2.74	6.689

Experimental setup

In this work, an experimental environment with the following environmental factors is built using the deep learning framework TensorFlow: CPU, Intel i7-6700 3.4GHz; GPU, Nvidia GTX 1060 graphics card; 8GB of PC memory; Windows 10 64-bit operating system; Python 3.6. More data features are present in datasets, and these features have differing degrees of impact on the anticipated value of dissolved oxygen. The most pertinent data that can aid in the model's ability to make correct predictions is selected using feature extraction and recognition in order to show the influence of input features on the model for predicting water quality. The proposed network layers listed in table 2. The degree to which features influence one another is now determined using Spearman's Rank correlation analysis.

Table-2 indicates details about layers in proposed network

Layer	**value**
Number of parameters	59
Number of hidden units	6
Number of inputs	16
Conv+ReLU	4
Max-pooling	7

Result and discussion

The proposed Improved Back Propagation Neural Network with Stochastic Gradient Descent (Imp BPNN SGD) is evaluated using metrics such as Root Mean Square Error (RMSE), Relative Absolute Error (RAE), Mean Absolute Error (MAE), and Decision Coefficient (D^2). These metrics are used to evaluate three baseline techniques, including Recurrent Neural Network (RNN), Deep Long Short Term Memory (DeepLSTM) and Artificial Neural Network (ANN). Where D2 refers for fit optimization, representing the correlation between the two random variables, and RMSE and MAE may be used to calculate the error between the model's predicted value and the actual value; the minimum number, the more precise the results, as seen in the given description.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - q_i)^2} \qquad (11)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - q_i| \qquad (12)$$

$$\text{RAE} = \frac{\sum_{i=1}^{n}(pi-Ai)^2}{\sum_{i=1}^{n} Ai} \qquad (13)$$

$$D^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-q_i)\text{^}2}{\sum_{i=1}^{n}(q_i-q'_i)\text{^}2} \qquad (14)$$

Where $\boldsymbol{y_i}$ indicates the forecast value, $\boldsymbol{q_i}$ indicates the real value, n is the number of test sample, and $\boldsymbol{q'_i}$ indicates average value of the actual sequence.

The analysis of RMSE values of proposed methods and existing methods are shown in Table 3.

Table-3 analysis of RMSE

Number of days	DeepLSTM	RNN	ANN	Imp_BPNN_SGD
10	57.9	46.6	67.3	31.3
20	58	48	68	30
30	54.3	49.3	65.9	33.2
40	53.5	47	65	32
50	55	45.5	67.9	31.8

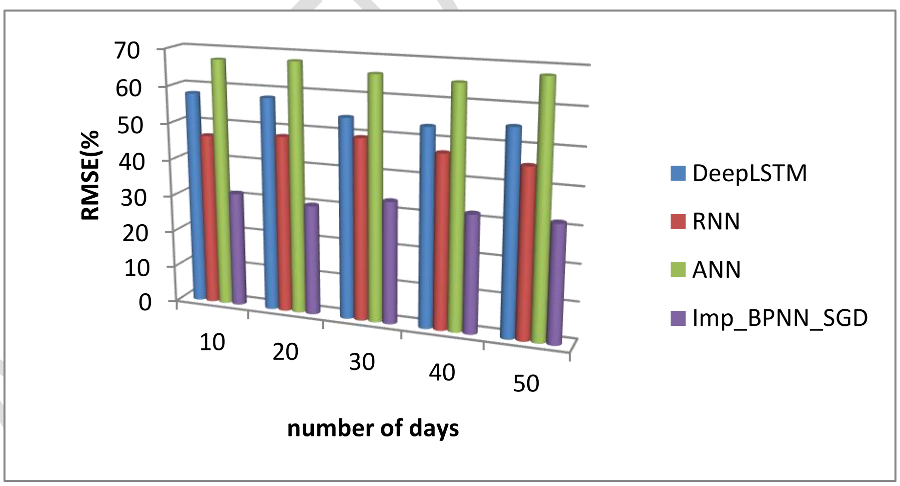


Figure-3 comparison of RMSE

Figure 3 shows the RMSE evaluation, where ,the suggested Imp_BPNN_SGD approach obtains 35.6% of RMSE, which is 23.9% lesser than the DeepLSTM method, 13.3% lesser than RNN and 32.6% lesser than ANN. The analysis of MAE values of proposed methods and existing methods are shown in Table 4.

Table-4 analysis of MAE

Number of days	DeepLSTM	RNN	ANN	Imp_BPNN_SGD
10	52	48	64	29
20	59	42	62	32
30	53	46	67	36
40	49	41	68	28
50	45	39	65	26

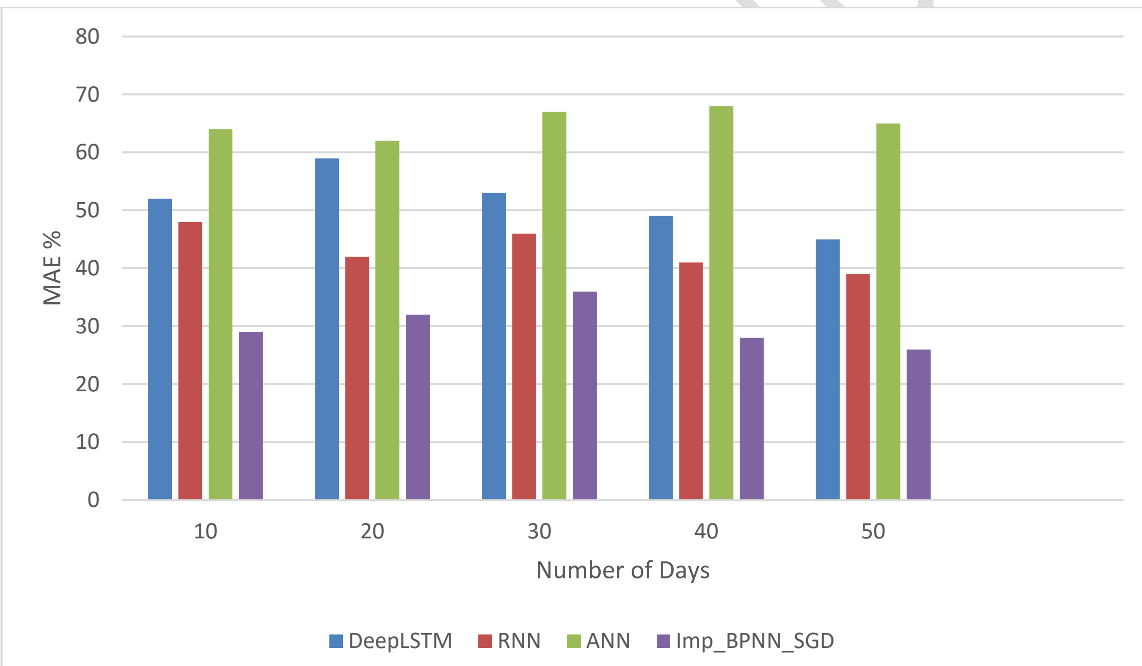


Figure-4 comparison of MAE

Figure 4 shows the MAE evaluation, where, the suggested Imp_BPNN_SGD approach obtains 21.6% of MAE, which is 26.2% lesser than the DeepLSTM method, 46.3% lesser than RNN and 33.3% lesser than ANN. The analysis of RAE values of proposed methods and existing methods are shown in Table 5.

Table-5 analysis of RAE

Number of days	DeepLSTM	RNN	ANN	Imp_BPNN_SGD
10	55	46	62	24
20	57	45	65	33
30	51	49	64	31
40	47	40	63	25
50	43	37	61	28

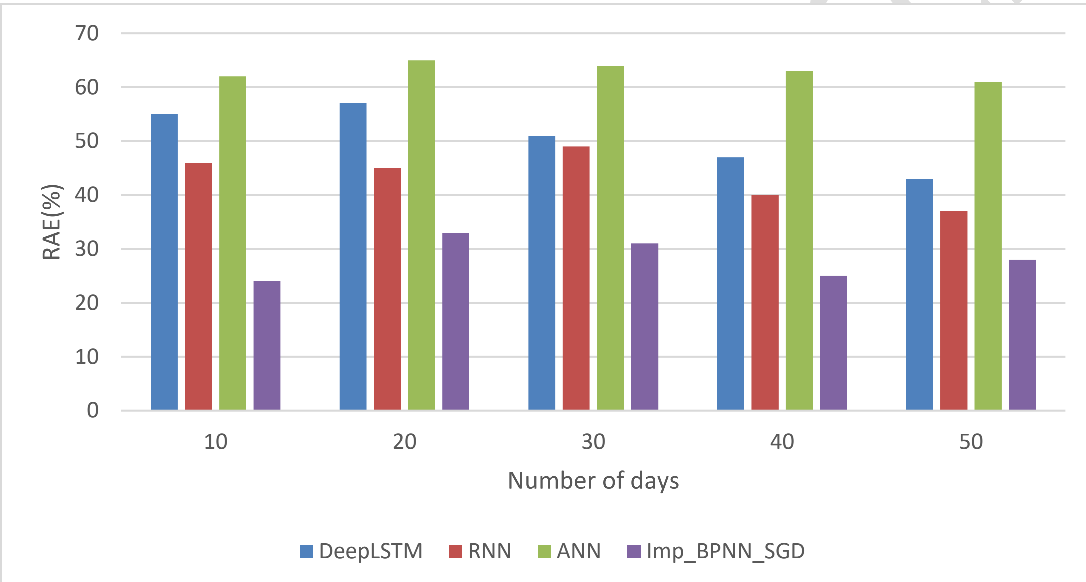


Figure-5 comparison of RAE

Figure 5 shows the RAE evaluation, where, the suggested Imp_BPNN_SGD approach obtains 19.6% of RAE, which is 37.8% lesser than the DeepLSTM method, 36.8% lesser than RNN and 42.2% lesser than ANN. The analysis of D^2 values of proposed methods and existing methods are shown in Table 6.

Table-6 analysis of D^2

Number of days	DeepLSTM	RNN	ANN	Imp_BPNN_SGD
10	65	82	73	97.5
20	68	80	76	95.7
30	62	78	71	96
40	59	79	82	96.9
50	72	85	86	94

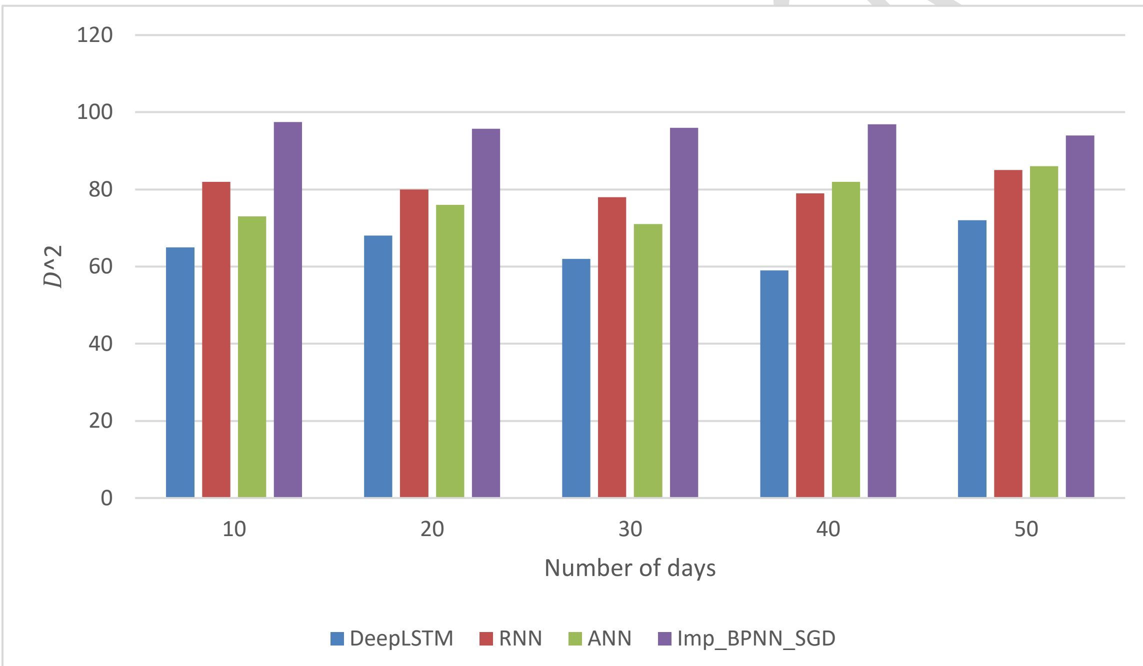


Figure-6 comparison of D^2

Figure 6 shows the D^2 evaluation, where, the suggested Imp_BPNN_SGD approach obtains 97.2% of D^2, which is 20.6% lesser than the DeepLSTM method, 16.2% lesser than RNN and 21.2% lesser than ANN.

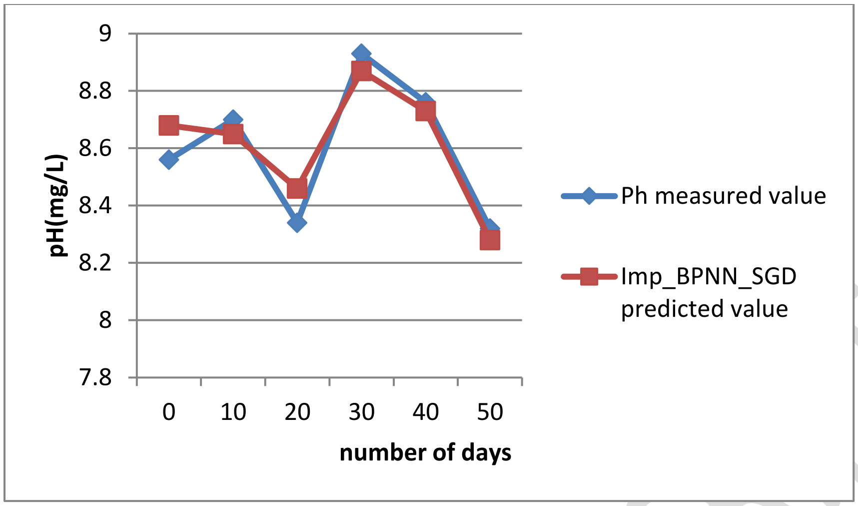


Figure-7 analysis of pH level in water using Imp_BPNN_SGD

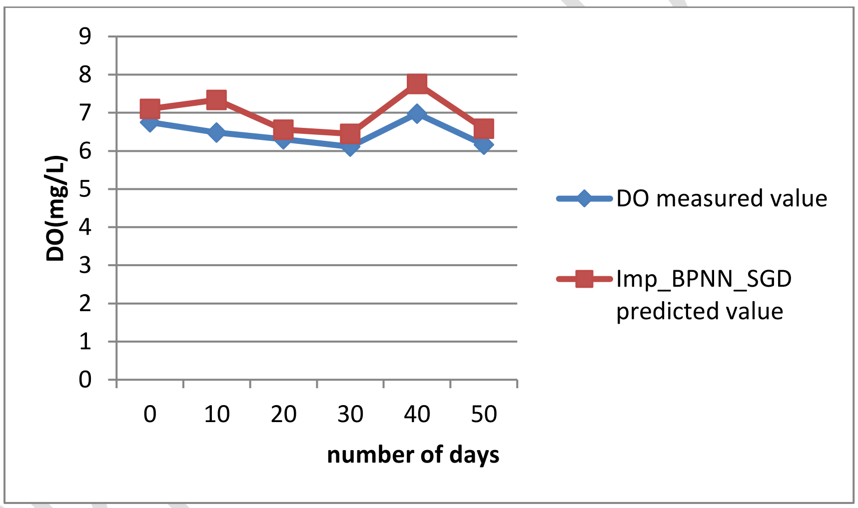


Figure-8 analysis of pH level in water using Imp_BPNN_SGD

In figures 7 and 8, the Imp BPNN SGD model's strong prediction abilities are further demonstrated by the five water quality parameters whose predicted values closely matched their measured values.

The outcome demonstrated that Imp BPNN SGD's prediction accuracy for measured data on coastal lagoon water was still high, demonstrating that Imp BPNN SGD's prediction performance was outstanding for application to actual prediction. The Imp BPNN SGD prediction will be helpful for aquaculture system monitoring and prognosis as well. Additionally, short-term forecasting is advised for improving forecast outcomes. The below table 7 shows the overall performance analysis for proposed method with existing methods.

Table 7. Comparison Overall performance Analysis

Parameters	**DeepLSTM**	**RNN**	**ANN**	**Imp_BPNN_SGD**
RMSE (%)	59.5	48.9	67.8	35.6
MAE (%)	47.8	67.9	54.9	21.6
RAE (%)	57.4	56.4	61.8	19.6
D^2 (%)	67.8	81.4	76.5	97.2

Conclusion

In this assessment, hourly groundwater levels observed at a coastal unrestrained aquifer in the Lagoon of Pulkit, India, are extended and simulated using an Improved Back Propagation Neural Network with Stochastic Gradient Descent (Imp BPNN SGD). First, utilising historical observed pore water data combined with external inputs, the constructed Imp BPNN SGD was trained to make one-step ahead predictions. After training, simulations were created by substituting run-time results from the Imp BPNN SGD with previously recorded data. In this method, the research evaluated Imp BPNN SGD's capacity to provide precise groundwater level simulations over an extended period of time. Additionally, it is observed that the constructed network provides effective performance than the baseline model which was used for comparison. In reality we can apply this research in all coastal regions. In a subsequent work, we plan to compare the Imp BPNN SGD results with a currently under construction numerical model of the aquifer system.

Reference

1. Fang, X., Hao, C. L., Cheng, C. G. et.al, “Analysis of the pressures and a quantitative spatial evaluation of the effects of human activity on the lagoon ecosystem: a case study in Xincun and Li’an Lagoons,” Applied Ecology and Environmental Research, vol. 16, no.5, 2018, pp. 6253-6266.

2. Holon, F., Marre, V., Mouquet, N., Bockel, T., Descamp, P., & Deter, J., “A marine ecosystem's state of degrading is explained by a predictive model based on various coastal anthropogenic pressures”, Journal for management and conservation. Biological Conservation, 222, 2018, pp.125-135.
3. Katsuki, K., Seto, K., Tsujimoto, A., Takata, H., et.al., “Relationship between regional climate change and primary ecosystem characteristics in a lagoon undergoing anthropogenic eutrophication”, journal of Coastal and Shelf Science, no.222, 2019, pp. 205-213.
4. Basatnia, N., Hosseini, S. R., & Muniz, P, “Performance evaluation of ecological status indicators based on soft-bottom macroinvertebrates: a study along the shallow Gomishan lagoon”, Brazilian Journal of Oceanography, no. 63, pp.363-378.
5. Zhu, X., Guo, H., Huang, J. Xu, W., & Mai, Y., “An ensemble machine learning model based on remote sensing data for estimating water quality in coastal areas,” Journal of Environmental Management”, 323, 2022, pp. 116187.
6. Wu, S., Qi, J., Yan, Z., Lyu, F., Lin, T., Wang, Y., & Du, Z., “Remote sensing and a spatiotemporal deep learning model for spatial and temporal evaluations of nutrients and water quality in coastal areas”, International Journal of Applied Earth Observation and Geoinformation, 112, 2021, pp. 102897.
7. Hassan, N., and Mateo, C. S., “Employing satellite data, machine learning for improving water quality,” In IOP Conference Series: Earth and Environmental Science, Vol. 842, No. 1, 2021, pp. 012018.
8. Aslan, S., Zennaro, F.,., & Zhang, Critto, A, “Recurrent neural networks for complicated coastal lagoon settings water quality assessment: A case study on the Venice Lagoon”, 2022.
9. Du, Z., Qi, J., Más-López, S., Furlan, E F., & Liu, R, “A spatially weighted neural network based water quality assessment method for large-scale coastal areas”, Environmental Science & Technology, vo. 55, no.4, pp. 2553-2563.
10. García del Toro, E. M., Woo , L. F., García-Salgado, S., , M. I., & Quijano, M. Wu Á, (2022) “Utilizing Artificial Neural Networks to Predict the Amount of Dissolved Oxygen in Surface Water Discharged into Coastal Lagoon,” International Journal of Environmental Research and Public Health, vol.19, no.8, pp. 4531.

11. Taormina, R., Chau, K. W., & Sethi, R, “Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon”, Engineering Applications of Artificial Intelligence, vol. 25, no.8, 2012, pp. 1670-1676.
12. Basatnia, N., Varotsos, S. A., Rodrigo-Comino, J, Dong, Khaledian, Y., Brevik, E. C., Aitkenhead-Peterson, J., & Natesan, U, “Evaluation of the international Gomishan Lagoon's water quality throughout time and space,” Environmental Monitoring and Assessment, vol. 190, no.5, 2018, pp. 1-17.
13. Krapivin, V. F., Hossein, C. A., & Nghia, B. Q, “a system for simulating the monitoring of lagoon water quality”, journal of Water, Air, & Soil Pollution, 2017, vol. 228, no. 10, 1-12.
14. Liu, Y., Wang, Y, F., Zou, R., Zhao, L., & He, B. “Assessment of potential water diversion scenarios' effects on lake eutrophication quantitatively,” Science of the total Environment, no. 468, pp. 219-227.
15. Melesse, A. M., Üstün, B., & Zhang, K, “Modeling coastal eutrophication at Florida Bay using neural networks’, Journal of Coastal Research, vol. 14, no. 3, 2008, pp. 90-196.
16. Taner, M. Ü.,. Krishnaswamy, J, & Erdinçler, A, “A straightforward instrument for evaluating the water quality in systems with polluted lagoons: A case study for Küçükçekmece Lagoon”, Turkey. Ecological Indicators, vol. 11, no. 2, 2011, pp.749-756.
17. Jimeno-Sáez, P., Sheng, H Senent-Aparicio, J., Cecilia, J. M., & Pérez-Sánchez, J, “Modeling eutrophication with machine learning algorithms”, International Journal of Environmental Research, Vol. 17, no. 4, 2020, pp.1189.