Vertical Electric Sounding and Multi-Layered Deming Perceptive Classifier for Ground Water Identification and Quality Prediction Analysis

Raghuveer Narsing*, K. Karthikeyan

Department of Civil Engineering, Annamalai University, Chidambaram, Tamilnadu, India

Received August 4, 2021; Revised October 13, 2021; Accepted October 27, 2021

Cite This Paper in the following Citation Styles

(a): [1] Raghuveer Narsing, K. Karthikeyan, "Vertical Electric Sounding and Multi-Layered Deming Perceptive Classifier for Ground Water Identification and Quality Prediction Analysis," Civil Engineering and Architecture, Vol. 9, No. 7, pp. 2249-2260, 2021. DOI: 10.13189/cea.2021.090712.

(b): Raghuveer Narsing, K. Karthikeyan (2021). Vertical Electric Sounding and Multi-Layered Deming Perceptive Classifier for Ground Water Identification and Quality Prediction Analysis. Civil Engineering and Architecture, 9(7), 2249-2260. DOI: 10.13189/cea.2021.090712.

Copyright©2021 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Water quality prediction is a critical problem because of vigorous impact on natural surroundings and human life. Groundwater crisis has become a global concern as their vulnerability increased with greater frequency and magnitude. This study proposes a Vertical Electric Sounding Ground Water Identification based Multi-Layered Perceptive Classifier (VESGWI-MLPC) model introduced to address prediction accuracy. The main objective is to perform groundwater quality prediction analytics with lesser time consumption and higher accuracy. This model performs two processes, namely groundwater identification and quality prediction analytics. Groundwater identification is carried out using a vertical electric sounding process. The vertical electric sounding process is a cost-effective resistivity survey process. Current is moved through the subsurface from one current electrode to the other and the current move is recorded for groundwater identification. Groundwater quality prediction analytics is carried out using a multi-layered perceptive classification process. In order to perform groundwater quality prediction analytics with higher accuracy, multi-layered perceptive is employed. Multi-layered perceptive classification process contains three layers for groundwater quality prediction analytics. In input layer, amount of information is collected as input. After that, parameters of the input data like temperature, pH, turbidity, salinity, nitrates, and phosphates are analyzed through perceptron in a hidden layer for performing groundwater quality prediction analysis. At last, output layer classifies and displays data. Thus, the groundwater quality prediction analytics performance gets increased in accuracy and time. Simulation results are considered by different factors namely prediction accuracy, error rate, and prediction time by amount of data. VESGWI-MLPC model enhances prediction accuracy and reduces prediction time and error rate compared with three state-of-the-art methods.

Keywords Water Quality, Vertical Electric Sounding, Groundwater Quality Prediction, Multi-Layered Perceptive Classification, Cost-Effective Resistivity

1. Introduction

Water pollution is described as any chemical or physical variations in water detrimental to the living organizations. Water pollution has become a demanding issue all over the globe causing useful water scarcity. The hybrid multivariate prediction model was introduced in [1] for seawater quality estimation of PCA and RVM. PCA is used for reduction of dimension. RVM was introduced to forecast oxygen and pH. However, accuracy is not enhanced with hybrid multivariate prediction scheme. A hybrid model is termed process-based watershed model.
and ANN is introduced in [2] for increasing prediction of water quality in unmonitored watersheds. Designed models were employed for developing calibration and validation process. However, error rate is not reduced by the hybrid model. The decision tree machine learning technique was employed in [3] to forecast the WQI for Klang River within water quality. Klang River was polluted river to analyse prediction and accuracy depending on different scenarios with water quality parameters. Though the classification accuracy is enhanced, time consumption is not reduced. Meta heuristic regression scheme is designed in [4] for estimating constituent scenarios. An ensemble ANN model depending on tiering method is more perfect than ensemble method and hybrid Meta heuristic regression schemes. But, computational cost was not reduced by the Meta heuristic regression model. A hybrid statistic scheme is introduced in [5] for joining ANN by GA to enhance DWTP performance in China. The HANN model was optimized through the lowest mean squared error (MSE). Though the error rate was reduced, time complexity was reduced by the hybrid statistic model. The machine learning (ML) technique was designed in [6] for prediction trend and magnitude of algal growth. SVM was appropriate for detecting finest scheme. ML technique guaranteed robustness to study difficulty among algal dynamics and coastal ecological variables. But, the time consumption was not reduced by the ML technique. FFNN framework was introduced in [7] for tuning hyper-parameters and GEP by PSO. However, specificity level was not reduced by the FFNN framework. An ensemble machine learning approach was designed in [8] for model stacking with a two-layered learning structure. The machine learning model considered input features to produce the final prediction. However, accuracy is not enhanced with ensemble machine learning approach. A hybrid model was introduced in [9] depending on LSTM based encoder-decoder neural network and Savitzky-Golay filter. Savitzky-Golay removed noise in time series of water quality. However, prediction time is not reduced with hybrid scheme. Water quality prediction scheme is developed in [10] with regression method. WQI was determined with arithmetic index technique. PCA was carried out with dominant WQI parameters. But, the specificity level was not reduced by the water quality prediction model.

Water quality monitoring has large significance for the stable operation of human society. Many studies employ the measured data from the monitoring station. A novel two-dimensional prediction of water quality sub-scheme is introduced in [11] for integrating two-dimensional water quality and DEM via GIS platform and FORTRAN language. A new hybrid decision-based machine learning technique was introduced in [12] to achieve short-term prediction of water quality results. An extreme gradient boosting (XGBoost) and RF introduced data denoising technique. A new RNN model was introduced in [13] to handle the long-term dependence for water quality prediction. A new method was introduced in [14] for prediction of water quality with higher accuracy. However, computational complexity is not minimized. WGP method is introduced in [15] to increase biochemical oxygen demand. Shannon entropy is employed for finding finest input mixture of WGP. But, accuracy is enhanced, computational cost is not reduced.

The correlation was determined in [16] with water quality parameters in water reservoir. But, the time consumption was not reduced by correlation measure. An effective data-driven model was designed in [17] for observing water quality variation and presented warning consistent with historical observation data. However, water quality accuracy is not enhanced. A hybrid model based on SAE and LSTM was introduced in [18] to increase accuracy of oxygen in aquaculture. The hidden layer data comprised deep latent features to improve the prediction accuracy. However, accuracy is increased complexity and is not reduced with SAE. A prediction of water quality method depending on LSTM was introduced in [19] to forecast pH and water temperature. But, the time consumption for water quality prediction is not reduced. Water quality prediction method was designed in [20] for regression and classification. SARIMANN model was developed in [22] to handle the high fluctuation series. But, the dry season performance was poor. Water Assessment Model (WEAP) was introduced in [21] for forecasting the water demand. However, the prediction accuracy was not enhanced.

The issues identified from the above review are higher time consumption, lesser accuracy, lesser error rate, higher computational cost, higher complexity, and so on. These problems are addressed by introducing a new model termed the VESGWI-MLPC model. The major contributions of the proposed VESGWI-MLPC model are summarized in the following lines:

- VESGWI-MLPC model is introduced to perform the groundwater quality prediction analytics with lesser time consumption and higher accuracy. VESGWI-MLPC model performs groundwater identification and quality prediction analytics.
- Groundwater identification is carried out using a vertical electric sounding process where the Current is moved through the subsurface from one current electrode to the other and the current move is recorded
- Groundwater quality prediction analytics is carried out using a multi-layered perceptive classification process. Multi-layered perceptive is employed using three layers to perform the groundwater quality prediction analytics with higher accuracy.
- This in turn helps to perform that the groundwater quality prediction analytics performance gets increased in terms of accuracy and time.
The paper is organized as follows. The description of the materials and methods is explained in section 2. Section 3 and 4 explain the result and discussion of proposed and existing methods. Section 5 concludes the paper.

2. Materials and Methods

Water is an essential requirement in daily life, which is used for agriculture, and stock performance. In order to achieve community, water resources were required. The groundwater is the national treasure facing the lack of surface water. Groundwater has become reliable supply of water source in every climatic region because of inherent qualities. Water pollution is the most important worldwide cause of death and disease. In addition, water pollution is a significant cause of death every day. The water quality is identified depending on the physical, chemical, and biological parameters. The water quality is identified through expensive time-consuming processes. A VESGWI-MLPC model is introduced to perform the quality prediction and to identify whether groundwater is suitable for drinking purposes. The proposed VESGWI-MLPC model used the vertical electrical sounding and multilayer perceptron to predict the water quality accurately.

Figure 1. Flow process of VESGWI-MLPC model

Figure 1 explains VESGWI-MLPC model for water quality prediction by better accuracy. Water quality dataset is measured for analyzing groundwater identification and quality prediction. At first, the data are collected and stored in the input database. The groundwater identification is carried out using Vertical Electrical Sounding. After finding the groundwater, the quality level is determined using Multilayer Perceptron Classifier depending on the physical, chemical, and biological parameters like temperature, pH, nitrates, and so on. The two processes of the VESGWI-MLPC model are explained briefly in the following subsections.

2.1. Vertical Electrical Sounding Groundwater Identification

VES was the geophysical process employed for groundwater identification in the geological medium. VES method is introduced depending on estimation of electrical conductivity or medium resistivity. VES estimation is carried out through the voltage measurement of the electric field induced by distant grounded electrodes. VES determined the resistivity variation with the depth and provide data with vertical distribution. From Schlumberger array, VES is performed. By small partition, possible electrodes are positioned in permanent position. Current electrodes are positioned on external side of possible electrodes. Behind resistivity capacity, current electrodes are moved from centre of array. Position of current electrodes was distributed with 10 positions per decade. The potential electrode distance improved for ensure calculates voltage of noise level and detection level. Depend on the principle of Vertical Electrical Sounding which allows current into ground by current electrodes and calculating potential drop among possible electrodes. By enhancing electrodes spacing, current penetrates the ground.

VES method is used for studying the subsurface layer properties and groundwater potential. The subsurface characterization is determined on resistivity values by depth. A resistivity value denotes occurrence of rock mass features. In order to determine depth-wise layer resistivity and breadth, Vertical electrical sounding is employed. Electrical resistivity capacity needs four-electrode plan. The two electrodes are used for inserting current within ground and other two are used to determine potential. Schlumberger configuration is used for VES data collection. Schlumberger electrode array is used in electrical prospecting for VES quantitative interpretation. Four electrodes were maintained at distance symmetrical to central point. The direct current (I) is inserted within ground via outer pair of electrodes and designed as current into ground is calculated over inner pair of electrodes.

Figure 2. Diagrammatic Representation of Schlumberger Electrode Array

Figure 2 illustrates a diagrammatic representation of the Schlumberger Electrode Array. ‘0’ represents the centre of the array ‘A’ and ‘B’ are the current electrodes where ‘M’ and ‘N’ are potential electrodes. The distance between M and N is ‘b’ small compared to ‘a’. It is formulated as,

$$\nabla V = \frac{I \rho_a}{2\pi} \left( \frac{1}{r_1} - \frac{1}{r_2} \right)$$  \hspace{1cm} (1)

In the case of potential electrodes difference ‘$\nabla V$’ among M and N is written in terms of notations. It is symbolized as,

$$\nabla V = \frac{I \rho_{as}}{2\pi} \left( \frac{1}{a - b} - \frac{1}{a + b} - \frac{1}{a + \frac{b}{2}} + \frac{1}{a - \frac{b}{2}} \right)$$ \hspace{1cm} (2)
The terms in the bracket are denoted by G. ‘b’ is small for the Schlumberger array. It is simplified as,

\[
G = \frac{2b}{a^2 - b^2} \tag{3}
\]

\[
G = \frac{2b}{k^2} \tag{4}
\]

Substituting equation (4) in (2) provides the relationship between potential ‘ΔV’ and apparent resistivity measured in the Schlumberger array. It is given as,

\[
\nabla V = \frac{I\rho_{as}}{\pi} \left(\frac{b}{a^2}\right) \tag{5}
\]

The apparent resistivity (\(\rho_{as}\)) of inhomogeneous ground is determined by (\(\nabla V\)) for current input by geometric factor (\(K\)) which is calculated by,

\[
\rho_{as} = K \frac{\nabla V}{I} \tag{6}
\]

From (6), ‘\(K\)’ denotes geometric factor on electrode partition. In order to establish vertical variation among electrical resistance and potential field developed by current, VES is employed in electrical resistivity review. The designed method included electric current induced within ground via two implanted electrodes and determined difference among two electrodes. Current electricity is a direct current presented by dry cell. Resistivity was determined by quantity of induced currents. Earth resistivity is established through standardized lithology and geological construction. In order to establish vertical variation, graph of resistivity to current electrode distance is employed. Parameter influenced groundwater resource estimation comprised the aquifer thickness and size in aquifer materials. The algorithmic process of vertical electrical sounding groundwater identification is shown below

```java
// Algorithm 1: Vertical Electrical Sounding Groundwater Identification
Input: Dataset, Station
Output: Identify the groundwater
Begin
for each station,
Perform refractor velocity analysis ‘\(\nabla V\)’
Calculate apparent resistivity ‘\(\rho_{as}\)’
Identify groundwater in the location of the particular station
End for
End
```

Algorithm 1 explains vertical electrical sounding groundwater identification. Let us consider that location in a particular station identify the groundwater. The optimal distance is determined based on refractor velocity analysis and apparent resistivity. Based on two-parameter calculations, the groundwater around the river delta region is recognized.

2.2. Multi-Layered Perceptive Classifier Based Predictive Analytics

Multilayer perceptron is a feed forward ANN which develops output via inputs. MLP is a family of machine learning of multiple layers for learning the given input. MLP is classified with many layers of input nodes linked when directed graph among input and output layers. MLP employs back propagation for education network. The proposed VESGWI-MLPC method uses multilayer perceptron to improve the performance of groundwater quality prediction analytics with minimal time consumption. MLP includes input, output by one hidden layer of non-linearly-activating nodes. Every node is completely associated with other consecutive layers. Every layer performed for executing final output. Multi-layer perceptron is illustrated in figure 3.

![Figure 3. Structure of multilayer perceptron](image)

Figure 3 illustrates the structural diagram of the multilayer perceptron. The deep structure comprises the neurons that are connected as one layer to others with help of consecutive layers. Input as one layer is associated to others in forward direction by dissimilar adjustable weights. Input layer is an amount of data ‘\(d_i \in d_1, d_2, d_3, ..., d_n\)’ and the activity of neurons in input layer ‘\(I(t)\)’ at time ‘\(t\)’ is shown as,

\[
I(t) = \sum_{i=1}^{n} d_i * w_{initial} + bias \tag{7}
\]

From (7), the input layers combines the input data ‘\(d_i\)’ with initial weight ‘\(w_{initial}\)’. The input is sent into the hidden layer. VESGWI-MLPC model uses the Deming regression analysis for performing groundwater identification and quality prediction analytics. VESGWI-MLPC model assumes that the available data (\(d_i, y_i\)) are measured observation of true values (\(d_i^*, y_i^*\)) based on the parameters of the input data like temperature, pH, turbidity, salinity, nitrates, and phosphates lie on the regression line as,
**Algorithm 2: Multilayer Perceptive Classifier Based Predictive Analytics**

Input: Number of data \( d_i \in d_1, d_2, d_3, \ldots, d_n \)

Output: Increase groundwater quality prediction

Begin

Collect the number of data \( d_i \in d_1, d_2, d_3, \ldots, d_n \) into the first hidden layer

For each data

Perform Deming regression analysis with parameters of input data

The hidden layer result is sent to the output layer

Update weights between layers

Output layer displays the result

End for

End

The algorithmic process of a multilayer perceptive classifier is described to attain accuracy and minimum error. Multilayer perceptive classifier collects the input data and gives to hidden layer. Deming regression is performed to analyze input data for performing groundwater quality prediction. The hidden result is transmitted to output layer. At last, exact output is shown from output layer. In this way, groundwater quality prediction has better accuracy and reduced time consumption.

### 3. Results

Simulation analysis of the proposed VESGWI-MLPC method and two existing methods, namely PCA-RVM [1] and process-based watershed model-ANN [2], is implemented using MATLAB simulator. In order to conduct experiment, two datasets are used, namely, the Indian water quality dataset (D1) and water samples gathered from the river bed of ANDHRA PRADESH state (D2). Depending on two datasets, the quality of water level gets examined for irrigation and drinking purposes. Indian water quality dataset is collected from https://www.kaggle.com/anbarivan/indian-water-quality-data. The dataset has 1991 instances and 12 attributes (i.e., columns). The dataset comprised the historical information on pollution levels and temperature across different states in India. For experimental consideration, the number of data points considered ranges from 100 to 1000. Correspondingly, the water quality parameters were collected from the river bed of ANDHRA PRADESH state and experiment with data in the range from 10 to 100. The 12 attributes and the descriptions are listed given below.

| S. No | Attributes                  | Description                     |
|-------|-----------------------------|---------------------------------|
| 1     | Station code                | Station code of every listed places |
| 2     | LOCATIONS                   | Address of the place where the data acquired from |
| 3     | STATE                       | Name of the state where the values are measured |
| 4     | temp                        | Temperature in °C               |
| 5     | D.O. (mg/l)                 | Dissolved oxygen in (mg/l)      |
| 6     | PH                          | Potential of Hydrogen           |
| 7     | CONDUCTIVITY                | Water can conduct an electrical current |
| 8     | B.O.D. (mg/l)               | Biochemical Oxygen Demand       |
| 9     | NITRATENANN + NITRITENANN (mg/l) | Chemical element            |
| 10    | FECAL COLIFORM              | Fecal coliform bacteria         |
| 11    | TOTAL COLIFORM              | Total coliforms include bacteria that are found in water |
| 12    | year                        | Year of data collected          |

\[ d_i = d_i^* + \rho_i \] \hspace{2cm} (8)

\[ y_i = y_i^* + \delta_i \] \hspace{2cm} (9)

From (8) and (9), \( \rho_i \) and \( \delta_i \) are independent and the ratio of their variances is considered to be known. For identifying the intercept \( \varphi_0 \) and the slope \( \varphi_1 \), it is formulated as,

\[ \hat{y}_i = \varphi_0 + \varphi_1 \delta_i \] \hspace{2cm} (10)

From (10), \( \delta_i \) and \( \hat{y}_i \) are the estimates of true values of \( y_i^* \) and \( d_i^* \), respectively. By using equations, the VESGWI-MLPC model identifies the best fit line for each data to perform the groundwater quality prediction. For every obtained result, the VESGWI-MLPC model minimizes sum of squared residuals and accurately predicts groundwater quality. The hidden layer result \( \hat{H}(t) \) is determined as,

\[ \hat{H}(t) = \sum_{i=1}^{n} d_i \times w_{ih} + w_{ho} \times \hat{Q}_i \] \hspace{2cm} (11)

From (11), \( w_{ih} \) represent the weight among input layer and hidden layer. Next, hidden layer are transferred from output layer. VESGWI-MLPC model transfer output from every input data. It is formulated as,

\[ O(t) = w_{ho} \times \hat{H}(t) \] \hspace{2cm} (12)

From equation (12), \( O(t) \) represent output layer. \( w_{ho} \) denotes weight allocated among hidden layer and output layer. Thus, data is classified as five categorize (i.e., drinking, irrigation, industries, domestic consumption, and water bodies) for performing groundwater quality prediction by better accuracy and reduced time consumption. Multilayer perceptive classifier based predictive analytics are explained below.
4. Discussion

The experimental results of proposed VESGWI-MLPC method and existing methods namely PCA-RVM [1], process-based watershed model-ANN [2] are described and examined. The comparison of the three methods is carried out with the help of a table and graph. The analysis of various techniques is presented in the following subsection. The performance of the VESGWI-MLPC model is measured in terms of three parameters like prediction accuracy, error rate, and prediction time.

4.1. Impact of Prediction Accuracy

It is referred as proportion of amount of data points are predicted properly to amount of data points gathered via dataset. It is determined as,

\[ \text{PreAcc} = \frac{\text{Number of data points that are correctly predicted}}{\text{Number of data points}} \times 100 \]  

(13)

From (13), ‘PreAcc’ represent the prediction accuracy. It is computed in percentage (%).

Table 2(a) and (b) explains prediction accuracy with amount of data by two datasets. Prediction accuracy results indicate that the VESGWI-MLPC model presents enhanced results than the existing methods PCA-RVM [1] and process-based watershed model-ANN [2]. By using dataset 1, 100 data are taken. In the first iteration, 92 data are correctly predicted through the classification and the prediction accuracy is 92% using the VESGWI-MLPC model. By using two existing methods, namely, PCA-RVM [1] and process-based watershed model-ANN [2] the number of data correctly predicted are 86 and 87 and their prediction accuracy is 86% and 87%, respectively. Similarly, the three methods are used to achieve the prediction accuracy outcomes. For each method, ten different outcomes are observed. The proposed technique is compared with traditional methods. Chemical properties of data are measured in the ANDHRA PRADESH and hence it improves the prediction accuracy. The proposed technique is compared with two existing methods. Figure 4 (a) (b) illustrates the diagrammatic representation of prediction accuracy for two datasets.
Figure 4(a). Comparison of Prediction accuracy for three different methods such as VESGWI-MLPC model, PCA-RVM, and process-based watershed model-ANN using dataset1

Figure 4(b). Comparison of Prediction accuracy for three different methods such as VESGWI-MLPC model, PCA-RVM, and process-based watershed model-ANN using dataset2

Figure 4 (a) (b) illustrates the results of water quality prediction accuracy by amount of data. Numbers of input data are considered in horizontal axis and accuracy of various methods was examined on the vertical axis in terms of percentage. As illustrated in the graph, there are three colors of cones like blue, red, and green that represent the prediction accuracy of three methods namely the VESGWI-MLPC model, PCA-RVM [1], process-based watershed model-ANN [2], respectively. The graphical plot shows that the VESGWI-MLPC model outperforms well than the other two conventional techniques. This is due to the application of multilayer perceptron to examine the chemical properties like temperature, PH value, conductivity (µmhos/cm), Nitratenan Fluorides in the groundwater. Depending on the analysis, identified water is considered to be high quality for drinking. Or else, water in the regions is not suitable for drinking because of its highly polluted condition. Finally, the machine learning model accurately predicts the water quality in the region around ANDHRA PRADESH state in India. The average value represents that the prediction accuracy of the VESGWI-MLPC model is enhanced as 7% and 6% compared with existing PCA-RVM [1] and process-based watershed model-ANN [2] respectively in dataset 1. In addition, the prediction accuracy is improved as 9% compared with PCA-RVM [1] and 6% compared with process-based watershed model-ANN [2] in dataset 2.
4.2. Impact of Error rate

It is referred as proportion of amount of data predicted wrongly to amount of data collected from dataset. It is determined by, 

\[ E_{Rate} = \left( \frac{\text{Number of data incorrectly predicted}}{\text{Number of data}} \right) \times 100 \] (14)

From (14), ‘\( E_{Rate} \)’ represents the error rate. It is determined in percentage (%).

Table 3(a) and (b) illustrates the performance result of error rate on amount of data with two datasets. It is expressed on amount of data taken via input dataset. From table 2, error rate results signify that the VESGWI-MLPC model presents more improved results than the existing methods PCA-RVM [1] and process-based watershed model-ANN [2]. By using dataset 1, 100 data are taken. In the first iteration, 8 data are incorrectly predicted through classification and prediction accuracy is 8% using the VESGWI-MLPC model. With two existing methods, namely PCA-RVM [1] and process-based watershed model-ANN [2] the number of data incorrectly predicted is 14 and 13 and their error rate is 14% and 13% respectively. Similarly, the error rate results are examined for three methods. After attaining ten results, proposed method is evaluated with existing methods. Likewise, chemical properties of data from the ANDHRA PRADESH state were established. The average of ten results representing error rate of VESGWI-MLPC model is reduced by two existing methods. Figure 5 (a) (b) illustrates the diagrammatic representation of the error rate for two datasets.

| Number of data | VESGWI-MLPC model | PCA-RVM | Process based watershed model-ANN |
|----------------|-------------------|---------|----------------------------------|
| 100            | 8                 | 14      | 13                               |
| 200            | 7                 | 12      | 10                               |
| 300            | 6                 | 14      | 12                               |
| 400            | 10                | 15      | 11                               |
| 500            | 6                 | 14      | 12                               |
| 600            | 9                 | 15      | 11                               |
| 700            | 7                 | 13      | 11                               |
| 800            | 6                 | 12      | 9                                |
| 900            | 6                 | 11      | 8                                |
| 1000           | 5                 | 12      | 9                                |

Table 3(b). Comparison of Error rate for three different methods such as VESGWI-MLPC model, PCA-RVM, and process-based watershed model-ANN using the dataset (D2)

| Number of data | VESGWI-MLPC model | PCA-RVM | Process based watershed model-ANN |
|----------------|-------------------|---------|----------------------------------|
| 10             | 10                | 30      | 20                               |
| 20             | 5                 | 15      | 10                               |
| 30             | 10                | 17      | 13                               |
| 40             | 7                 | 12      | 10                               |
| 50             | 6                 | 12      | 10                               |
| 60             | 8                 | 13      | 12                               |
| 70             | 7                 | 14      | 11                               |
| 80             | 9                 | 15      | 14                               |
| 90             | 8                 | 13      | 12                               |
| 100            | 9                 | 14      | 12                               |

Figure 5(a). Comparison of error rate for three different methods such as VESGWI-MLPC model, PCA-RVM, and process-based watershed model-ANN using dataset1
Figure 5 (a) (b) illustrates the experimental results of error rate depending on amount of data. From graph, proposed VESGWI-MLPC model obtains a lesser error rate hence it reduces the error rate for both datasets. This is due to the application of a multilayer perceptron classifier. VES method was employed to study subsurface properties as well as groundwater potential. In order to estimate depth-wise layer resistivity as well as thickness, VES is performed. VESGWI-MLPC model uses the Deming regression analysis for performing the groundwater identification and quality prediction analytics. The multilayer perceptron classifier categorizes the accurate classification results with a minimum error rate. The average value represents that the error rate of the VESGWI-MLPC model is reduced by 47% and 33% when compared to existing PCA-RVM [1] and process-based watershed model-ANN [2] respectively in dataset 1. In addition, the error rate is reduced by 47% compared with PCA-RVM [1] and 36% compared with process-based watershed model ANN [2] in dataset 2.

4.3. Impact of Prediction Time

It is described as amount of time consumed for forecasting water quality on classification process. Water quality prediction time is determined as,

$$Pre_{time} = n \times time_{consumed \ for \ predicting \ one \ data}$$  \hspace{1cm} (15)

From (15), ‘$Pre_{Time}$’ represents the prediction time, ‘$n$’ symbolizes number of data collected via input dataset. It is calculated in milliseconds (ms).

Experimental analysis of water quality prediction time with amount of data via two datasets and three methods namely VESGWI-MLPC model, PCA-RVM [1], Process-based watershed model-ANN [2] are illustrated in table 4(a) (b). The observed table values indicate that the VESGWI-MLPC model reduced the time consumption for groundwater quality prediction. The experiment is conducted for 100 data from dataset 1, the time consumption of VESGWI-MLPC model for predicting water quality is 17ms whereas time consumption of prediction using existing PCA-RVM [1], Process-based watershed model-ANN [2] are 20ms and 19ms respectively. Likewise, additional results are attained with different input data.

| Number of data | VESGWI-MLPC model | PCA-RVM | Process-based watershed model-ANN |
|----------------|-------------------|--------|----------------------------------|
| 100            | 17                | 20     | 19                               |
| 200            | 18                | 24     | 20                               |
| 300            | 24                | 27     | 26                               |
| 400            | 24                | 30     | 28                               |
| 500            | 25                | 33     | 30                               |
| 600            | 27                | 35     | 33                               |
| 700            | 28                | 39     | 35                               |
| 800            | 28                | 42     | 36                               |
| 900            | 29                | 45     | 36                               |
| 1000           | 30                | 48     | 35                               |
Table 4(b). Comparison of Prediction time for three different methods such as VESGWI-MLPC model, PCA-RVM, and process-based watershed model-ANN using the dataset (D2)

| Number of data | VESGWI-MLPC model | PCA-RVM | Process based watershed model-ANN |
|----------------|-------------------|---------|-----------------------------------|
| 10             | 8                 | 11      | 10                                |
| 20             | 9                 | 12      | 10                                |
| 30             | 10                | 14      | 12                                |
| 40             | 12                | 15      | 14                                |
| 50             | 13                | 16      | 15                                |
| 60             | 13                | 17      | 15                                |
| 70             | 14                | 18      | 15                                |
| 80             | 14                | 19      | 16                                |
| 90             | 15                | 21      | 17                                |
| 100            | 16                | 22      | 18                                |

Figure 6(a) (b) illustrates the diagrammatic representation of prediction time for two datasets. The graphical representation of prediction time for two datasets shows the improvement in the prediction time with the increase in the amount of data. The VESGWI-MLPC model shows better prediction time compared to PCA-RVM and process-based watershed model-ANN.

Figure 6(a) (b) explains the graphical representation of water quality prediction time for the different numbers of data from two datasets. As illustrated in graphical representation, the number of data considered on the horizontal axis and different prediction time performance results are attained on the vertical axis. Prediction time of three methods is improved with enhancing amount of data. However, quality prediction time gets minimized for both two datasets using the VESGWI-MLPC model than the other two existing methods.

This is due to the application of vertical electrical sounding groundwater identification and quality prediction. The ensemble technique examines the training and testing value of the chemical properties of water. The water quality in selected regions is accurately predicted for drinking purposes with minimum time consumption. Prediction time is reduced as 5% compared with [1] and 12% compared with [2] respectively. Similarly in dataset 2, the average of ten results indicates that the overall quality prediction time gets reduced as 25% and 13% compared with existing methods.
5. Conclusions

In this paper, a Vertical Electric Sounding Ground Water Identification based Multi-Layered Perceptive Classifier (VESGWI-MLPC) model is introduced. VESGWI-MLPC model performs the groundwater quality prediction analysis with lesser time consumption and higher accuracy. The following conclusions and recommendations can be obtained from this study:

Experimental evaluation is evaluated for estimating VESGWI-MLPC model with two existing methods. VESGWI-MLPC model is performed as higher accuracy, lesser time consumption, and lesser error rate than conventional methods during groundwater quality prediction analysis.

The results of the study indicate that there is groundwater quality prediction analytics performance gets increased in terms of accuracy and time. In this study, Multi-layered perceptive classification process performs the groundwater quality prediction analytics. The parameters of input data like temperature, pH, turbidity, salinity, nitrates, and phosphates are examined through Deming regression for performing groundwater quality prediction analysis.

VESGWI-MLPC model increases the number of data points that is properly predicted as compared to existing works. VESGWI-MLPC model decreases the amount of time needed for predicting the water quality as compared to conventional works. Besides, the proposed VESGWI-MLPC model lessens the number of data points incorrectly predicted when compared to existing works. However, this proposed method, the VESGWI-MLPC model is suitable for water identification and quality prediction analysis and enhanced performance with improvement of accuracy and decreases prediction time compared with state-of-art methods.

REFERENCES

[1] Sun X., Wang X., Cai D., Li Z., Gao Y., Wang X, “Multivariate Seawater Quality Prediction Based on PCA-RVM Supported by Edge Computing Towards Smart Ocean”, IEEE Access, Vol.8, pp. 54506-54513, 2020. doi: 10.1109/ACCESS.2020.2981528

[2] Noori N, Kalin L, Isik S., “Water quality prediction using SWAT-ANN coupled approach”, Journal of Hydrology, Elsevier, Vol.590, 1-12, 2020. https://doi.org/10.1016/j.jhydrol.2015.08.066

[3] Ho J.Y., Afan H.A, El-Shafie A.H., Koting S.B, Mohd N.S., Jaafar W.Z.B., Hin L.S., Malek M.A., Ahmed A.N., Mohtar W.H.M.W., Elshorbagy A., El-Shafie A., “Towards a time and cost effective approach to water quality index class prediction”, Journal of Hydrology, Elsevier, Vol.575, 148-165, 2019. https://doi.org/10.1016/j.jhydrol.2019.05.016

[4] Chou J., Ho C., Hoang H., “Determining quality of water in reservoir using machine learning”, Ecological Informatics, Elsevier, Vol. 44, 57-75, 2018. https://doi.org/10.1016/j.ecoinf.2018.01.005

[5] Zhang Y., Gao X., Smith K., Inial G., Liu S., Conil L.B., Pan B, “Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network”, Water Research, Elsevier, Vol. 164, 1-15, 2019. https://doi.org/10.1016/j.watres.2019.114888
2260 Vertical Electric Sounding and Multi-Layered Deming Perceptive Classifier for Ground Water Identification and Quality Prediction Analysis

[6] Deng T., Chau K., Duan H., “Machine learning based marine water quality prediction for coastal hydro-environment management”, Journal of Environmental Management, Elsevier, Vol. 284, 1-18, 2021. https://doi.org/10.1016/j.jenvman.2021.112051

[7] Shah M.I., Javed M.F., Alqhtani A., Aldrees A., “Environmental assessment based surface water quality prediction using hyper-parameter optimized machine learning models based on consistent big data”, Process Safety and Environmental Protection, Elsevier, Vol.151, 324-340, 2021. https://doi.org/10.1016/j.psep.2021.05.026

[8] Wang L., Zhu Z., Sassoubre L., Yu G., Liao C., Hu Q., Wang Y., “Improving the robustness of beach water quality modelling using an ensemble machine learning approach, Science of The Total Environment”, Elsevier, Vol.765, 1-10, 2021. doi 10.1016/j.scitotenv.2020.142760

[9] Bi J., Lin Y., Dong Q., Yuan H., Zhou M.C., “Large-scale water quality prediction with integrated deep neural network”, Information Sciences, Elsevier, Vol.571, 191-205, 2021. https://doi.org/10.1016/j.ins.2021.04.057

[10] Khan S.I., Islam N., Uddin J., Islam S., Nasir M.K., “Water quality prediction and classification based on principal component regression and gradient boosting classifier approach”, Journal of King Saud University - Computer and Information Sciences, Elsevier, 1-18, 2021. https://doi.org/10.1016/j.jksuci.2021.06.003

[11] Ding X., Dong X., Hou B., Fan G., Zhang X., “Visual platform for water quality prediction and pre-warning of drinking water source area in the Three Gorges Reservoir Area”, Journal of Cleaner Production, Elsevier, Vol. 309, 1-12, 2021. https://doi.org/10.1016/j.jclepro.2021.127398

[12] Lu H., Ma X., “Hybrid decision tree-based machine learning models for short-term water quality prediction”, Chemosphere, Vol.249, 1-15, 2020. https://doi.org/10.1016/j.chemosphere.2020.126169

[13] Li L., Jiang P., Xu H., Lin G., Guo D., Wu H., “Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang River”, China, Environmental Science and Pollution Research, Springer, Vol. 26, 19879-19896, 2019. doi 10.1007/s11356-019-05116-y

[14] Zou Q, Xiong Q, Li Q, Yi H., Yu Y., Wu C., “A water quality prediction method based on the multi-time scale bidirectional long short-term memory network”, Environmental Science and Pollution Research, Springer, Vol.27, 16853-16864, 2020. doi:10.1007/s11356-020-08087-7

[15] Jafari H., Rajaei T., Kisi O., “Improved Water Quality Prediction with Hybrid Wavelet-Genetic Programming Model and Shamon Entropy”, Natural Resources Research, Vol.29, 3819-3840, 2020. doi:10.1007/s11053-020-09702-7

[16] Aghel B., Rezaei A., Mohadesi M., “Modelling and prediction of water quality parameters using a hybrid particle swarm optimization–neural fuzzy approach”, International Journal of Environmental Science and Technology, Springer, Vol.16, 4823-4832, 2019. doi:10.1007/s13762-018-1896-3

[17] Jin T., Cai S., Jiang D., Liu J., “A data-driven model for real-time water quality prediction and early warning by an integration method”, Environmental Science and Pollution Research, Vol.26, 30374-30385, 2019. doi: 10.1007/s11356-019-06049-2.

[18] Li Z, Peng F, Niu B, Li G., Wu J., Miao Z., “Water Quality Prediction Model Combining Sparse Auto-encoder and LSTM Network”, IFAC-PapersOnLine, Elsevier, Vol.51, 831-836, 2018. https://doi.org/10.1016/j.ifacol.2018.08.091

[19] Hu Z., Zhang Y., Zhao Y., Xie M, Zhong J., Tu Z., Liu J., “A Water Quality Prediction Method Based on the Deep LSTM Network Considering Correlation in Smart Mariculture”, Sensors, Vol.19, 14-20, 2019. https://doi.org/10.3390/s19061420

[20] Sheng L., Zhou J., Li X., Pan Y., Liu L., “Water quality prediction method based on preferred classification, IET Cyber-Physical Systems: Theory and Applications”, Vol.5, 176-180, 2020. DOI:10.1049/iet-cps.2019.0062

[21] Nualtong K., Pаниtпакул T., Khwanmuang P., Chinram R., Kirtsaeng S., "A Hybrid Seasonal Box Jenkins-ANN Approach for Water Level Forecasting in Thailand," Environment and Ecology Research, Vol. 9, No. 3, pp. 93 - 106, 2021. doi:10.13189/eer.2021.090301.

[22] Noon A.M., Ahmed H.G. I., Sultàıman S.O., "Assessment of Water Demand in Al-Anbar Province- Iraq," Environment and Ecology Research, Vol. 9, No.2, pp. 64 - 75, 2021. doi:10.13189/eer.2021.090203.