International Journal Of Engineering Research And Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 21, Issue 5 (May 2025), PP 182-195

Xg Boost Model Application in Spatio-Temporal Differences of Groundwater Quality in Different Geographical Location

Ikpa P. N., Osuagwu J. C., Nwakwasi Ikpa P. N., Uzoukwu C. S., Ugwuegbulam J. C., Obumseli P. C., Ezenkwa C. S., Nwokorobia G. C., Njoku K. O., Igbojiaku A.U., Onuoha O. I,, Osukalu E. J.
Civil Engineering Department,School Engineering and Engineering Technology,Federal University of Technology Owerri,Imo State, Nigeria

Abstract: *This study dealt with the XG boost model in analyzing the spatio-temporal variations of groundwater quality parameters from different hydrological formations of Imo State, Nigeria. Several iterative processes were conducted using this model, from the values generated from the processes as discussed below. On a monthly basis for a period of one year, 72 groundwater samples were collected from 6 hydrological formations of Imo State and were analyzed for 22 physiochemical parameters: pH, electrical conductivity, total dissolved solid, total suspended solid, total solid, dissolved Oxygen, Biochemical Oxygen Demand, (BOD), Chemical Oxygen Demand, Nitrate, Bicarbonate, (HCO_3)Potassium (K),Phosphate, (PO_4^{-3}), Iron, (Fe), total alkalinity, ($CaCO_3$) total Chloride, (Cl^-), Calcium hardness, total hardness, Magnesium hardness, Calcium (Ca),Magnesium (Mg),Sodium (Na) and Sulphate (SO_4^{-2}).The geological zones considered were the Benin Formation (BF), Ogwashi Asaba Formation (OAF), Nsukka Formation (NF), Alluvium Formation (AF), Imo Clay Shale Formation (ICSF), and False Bedded Sandstones Formation (FBSF). The average concentrations of total dissolved solids, chloride, nitrate, sulphate, total hardness and electric conductivity were higher in the dry season compared to the rainy season, while average concentrations potassium and bicarbonate were higher in wet season. The water quality index (WQI) was evaluated in accordance with WHO permissible standards for safe drinking water on a scale of 0 to 100. The WQI for dry season were 50.10, 24.98, 20.18, 35.79, 79.77 and 55.94 for BF, OAF, NF, AF, ICSF, and FBSF respectively while for rainy season, the WQI gotten were 35.04, 73.30, 27.54, 30.37, 86.98 and 108.95 for BF, OAF, NF, AF, ICSF and FBSF respectively. The results reveal that during dry season, groundwater samples from OAF and NF have excellent water quality, samples from BF, NF, and AF have good quality water and samples from ICSF have very poor water quality. The WQI obtained during the rainy season indicate that water samples from BF, NF and AF have good water quality for drinking and agricultural applications based on national and international indices and standards while the water samples from OAF were of poor water quality. The water quality from ICSF is very poor and the water quality from FBSF unsuitable for drinking purpose. This suggests that there is need for continuous monitoring and treatment for acidic and high nitrate water to mitigate future pollution and ensure sustainable use of the groundwater resource.*
XG Boost Model was used to model the dataset and a value of 142.829234 was obtained as the RMSE which subsequently decreased to 130.309532 at the final iteration.

Keywords: *XG Boost Model, Physiochemical parameters, Water quality, Spatio-temporal variability, Hydrological formations*

Date of Submission: 10-05-2025 Date of acceptance: 20-05-2025

I. Introduction

Groundwater pollution can have effects on poor drinking water quality, water supply interruption, degraded surface water systems, costly remedial action, the necessity for additional water sources, and/or possible health problems. Groundwater contaminated or degraded surface water could have negative impacts. Sundara et al. (2010) claim that groundwater has a spectrum of elements at different concentrations: gases, microbes, inorganic and organic compounds, etc. These concentrations create a concern and are regarded as undesirable contaminants when they exceed WHO drinking water recommendations (Amangabara & Ejenma 2012). Oladipo et al. (2014) claim that water pollution with trace metals can result from contaminated water seeping into the groundwater through rock and soil, as well as from prolonged exposure to intense sunlight, high temperatures, fragmentation, biological activity, etc., tend to bring bacteria or viruses into the water and water dissolves the minerals that are soluble in sedimentary rocks and soils. Thus, maintaining consumer safety and lowering the frequency of infectious diseases depend on constant observation of groundwater quality.

Many theoretical models exist for groundwater quality change both geographically and temporally. Fryar et al. (2000) looked at the spatial and temporal fluctuations in seepage between a contaminated aquifer and Ohio River tributaries. Hayashi and Rosenberry (2002) investigated how surface water hydrology and ecology responded to groundwater exchange.

Fryar et al. (2000) studied the temporal and spatial fluctuations in seepage between a contaminated aquifer and Ohio River tributaries. Hayashi and Rosenberry (2002) investigated how surface water hydrology and ecology might be affected by groundwater exchange. Allison (2005) investigated throughout time and space how groundwater discharge to streams changed. Mini et. al., 2014 investigated the temporal and geographical behavior of groundwater level in the coastal aquifers of Minjur in Tamilnadu, India using the GS+ and geostatistical modules of Arc GIS 9.3 software. They found that groundwater level exhibits notable spatial dependency. Dhar et al. (2008) examined the temporal variability of groundwater chemistry in shallow and deep aquifers in Araihazar Bangladesh and identified a link between aquifer age and mobility of Ions like As, Fe and so forth independent of the redox impact. Essien and Abasifreke (2004) investigated groundwater quality in boreholes located in the urbanized state capital of Uyo as well as four adjacent local government areas (LGAs) of Ibiono Ibom, Ikot Ekpene, Itu, and Nsit Ubium, all under the formation of coastal plain sands (CPS), in order to ascertain the spatial and temporal variability of groundwater quality and its fit with Nigerian Standards for Drinking Water Quality (NSDWQ). Their results suggested that the spread of urbanization could lead to pollution diffusion. Three types of boreholes: government-owned public boreholes, privately owned boreholes used for personal use, and individual-owned boreholes utilized for business usage were investigated by Agunwamba et al. (2000). Maintenance, a serious issue, could affect the quality of the groundwater released.

Objectives of Study: The major goal of this work is to simulate the Spatio-Temporal fluctuation of groundwater quality in several geological formations in Imo State. Other particular goals, though, include:
i) To evaluate groundwater quality by means of laboratory technique examination of a few chosen water quality criteria.
ii) To evaluate the findings against World Health Organisation (WHO 2017), FMEnv (2012) and BIS (2015) allowable limits.
iii.) Collecting samples during both wet (April to October) and dry season (November to March) can help one ascertain the effect of time on the chosen physio-chemical parameters.
iv) To ascertain and calculate the Water Quality Index (WQI) of some particular criteria.
v) Using XGBoost Model, to replicate the Spatio-temporal fluctuation of the water quality in several geological formations of Imo State.

II. Materials

The materials employed for this research work are:

Microsoft office package software, Google chrome and Mozilla firefox browser. In the list of hardware include, Intel Pentium Dell inspiron 5000, 4 GB RAM, HID Optical Mouse, HP Deskjet Ink Advantage 1515 printer, Tecno Pova Neo., Stop watch, Thermometer and pH meter, Atomic absorption spectrophotometer (AAS), 300 ml and 250 ml Amber DO and BOD bottles, Conductivity/TDS Meter, Spectrophotometer,Whatman Filter Paper, Pipettes and burettes, MnSO4 solution, Alkali-Iodide-Azide solution, K_2CrO_7 solution, $Ag_2SO_4 – H_2SO_4$ solution and Fe $(NH_2)_2$ $(SO_4)_2$.$6H_2O$ solution, Phenolphthalein indicator, P-nitrophenol, Ascorbic Acid and Sodium Acetate,Alkaline Phenol, Sodium Potassium Titrate, Sodium Hypochlorite and Brucine, Weighing scale, mercury in glass thermometer, Durham tubes, incubator, oven, and turbidity meter., Water bath, electrode colony counter etc.

III. Methodology

XG Boost Model:
The model was developed to predict the Spatio-temporal variability of groundwater quality in the various geological zones of Imo State.

- **Gradient Boosting Framework:**

The model application is an implementation of the gradient boosting algorithm, which builds models sequentially, with each new model focusing on correcting the errors of the previous ones.

Decision Trees as Base Learners:
The uses decision trees as its base learners, which are then combined to create a strong predictive model.

Regularization:
The model uses incorporates regularization techniques to prevent overfitting and improve the generalization ability of the model.

Parallel Processing:
The model is designed for efficient processing, utilizing parallel processing to speed up training, especially on large datasets.

Handling Missing Values:
The model t can handle missing data effectively, making it robust for real-world datasets.

Advantages of the Model:

High Accuracy:
The model often achieves state-of-the-art results in various machine learning competitions.

Speed and Efficiency:
Its optimized implementation and parallel processing capabilities make it very fast, especially for large datasets.

Scalability:
XGBoost can handle very large datasets, making it suitable for real-world applications.

Flexibility:
It can be used for both regression and classification problems, as well as ranking tasks.

Regularization:
The built-in regularization techniques help prevent overfitting and improve model generalization.

Disadvantages of The Model:

Overfitting:
The model can be prone to overfitting, especially on small datasets or with too many trees.

Interpretability:
While feature importance scores are available, the overall model can be challenging to interpret compared to simpler methods like linear regression or decision trees.

In summary, XGBoost is a powerful and versatile machine learning algorithm that is well-suited for a wide range of predictive modeling tasks, especially when speed, accuracy, and scalability are important.

In machine learning, different algorithms are often combined to get better and optimized results. The main goal is to minimize loss function for which, one of the famous algorithm is XGBoost (Extreme boosting) technique which works by building an ensemble of decision trees sequentially where each new tree corrects the errors made by the previous one. It uses advanced optimization techniques and regularization methods that reduce overfitting and improve model performance.

Learning Rate (eta): An important variable that modifies how much each tree contributes to the final prediction. While more trees are needed smaller values frequently result in more accurate models.
Max Depth: This parameter controls the depth of every tree, avoiding over-fitting and being essential to controlling the model's complexity.
Gamma: Based on the decrease in loss it determines when a node in the tree will split. The algorithm becomes more conservative with a higher gamma value, avoiding splits that don't decreases the loss. It helps in managing tree complexity.
Subsample: Manages the percentage of data that is sampled at random to grow each tree hence lowering variance and enhancing generalization. Setting it too low could result in underfitting.
Colsample Bytree: Establishes the percentage of features that will be sampled at random for growing each tree.
n_estimators: Specifies the number of boosting rounds.

lambda (L2 regularization term) and alpha (L1 regularization term): Control the strength of L2 and L1 regularization respectively. A higher value results in stronger regularization.
min_child_weight: Influences the tree structure by controlling the minimum amount of data required to create a new node.

scale_pos_weight: Useful in imbalanced class scenarios to control the balance of positive and negative weights.

The model is the leading model for working with standard tabular data (the type of data you store in Pandas Data Frames, as opposed to more exotic types of data like images and videos). It goes through cycles that repeatedly build new models and combines them into an **ensemble** model. It starts the cycle by calculating the errors for each observation in the dataset. it then builds a new model to predict those. Predictions are added from this error-predicting model to the "ensemble of models."

To make a prediction, the predictions from all previous models are added. It can use these predictions to calculate new errors, build the next model, and add it to the ensemble.There's one piece outside that cycle. It needs some base prediction to start the cycle. In practice, the initial predictions can be pretty naive. Even if it's predictions are wildly inaccurate, subsequent additions to the ensemble will address those errors.This process may sound complicated, but the code to use it is straightforward.

.

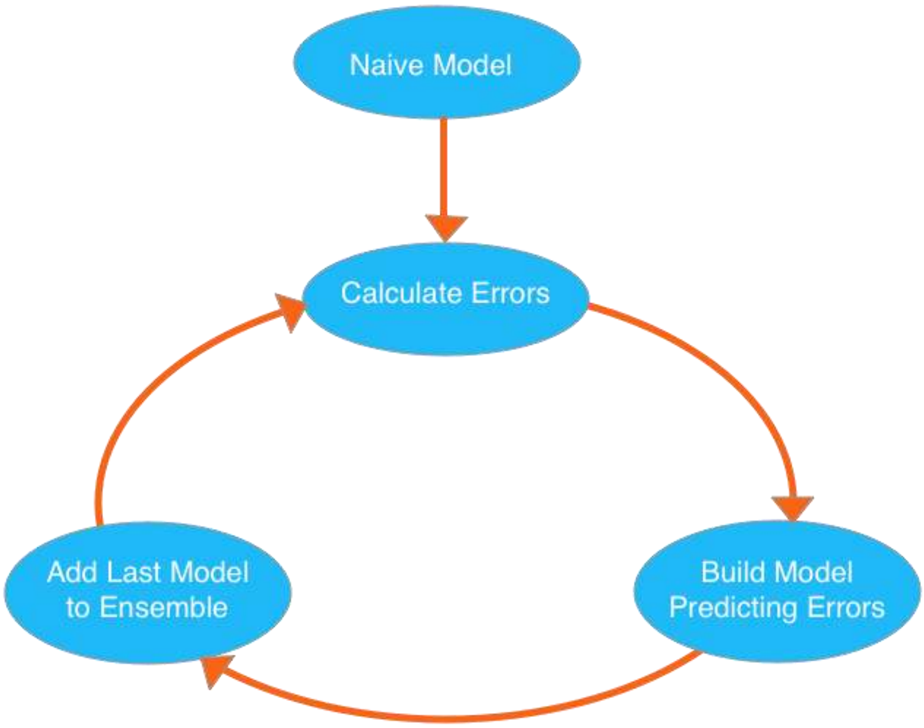

Figure 4.1: XGBOOST Model Cycle.

The XGBoost (XGB) model was trained using a 70:30 ratio to predict the variable Parameter_value based on two features: Season and Region. The model underwent 100 boosting iterations, adding sequential trees to refine predictions. Despite its simplicity, the model effectively captures interactions between the two features, as evidenced by the stabilized root mean square error (RMSE) of 130.3095 at the final iteration, indicating convergence and limited potential for improvement with further iterations.

XGBoost offers several advantages, including regularization techniques that prevent overfitting and its ability to model complex feature interactions, thus providing insights into how Season and Region influence the target variable. The decreasing RMSE from 142.8292 to 130.3095 suggests enhanced prediction accuracy throughout training. Hence, XGBoost is well-suited for regression tasks, combining speed and efficiency while laying a strong foundation for actionable data insights.

Hence, the XGBoost model based on the outcome can be expressed as:

$$\hat{y} = \sum_{k=1}^{100} f_k(x)$$

Where: $f_k(x)$is a decision tree defined by the parameters set within the training process of the explanatory variables, optimizing for squared error as the loss function and controlling complexity through the regularization term.$\hat{y}$ is the predicted response variable.

IV. Results

Table 4.1-3 Results of the Root mean Square error of the XGBoost model for 100 iteration

Iteration	RMSE
[1]	train-rmse:142.829234
[2]	train-rmse:136.904543
[3]	train-rmse:133.744519
[4]	train-rmse:132.089647
[5]	train-rmse:131.229497
[6]	train-rmse:130.784337
[7]	train-rmse:130.554444
[8]	train-rmse:130.435839
[9]	train-rmse:130.374677
[10]	train-rmse:130.343249
[11]	train-rmse:130.327315
[12]	train-rmse:130.318703
[13]	train-rmse:130.314261
[14]	train-rmse:130.311973
[15]	train-rmse:130.310793
[16]	train-rmse:130.310183
[17]	train-rmse:130.309868
[18]	train-rmse:130.309704
[19]	train-rmse:130.309621
[20]	train-rmse:130.309577
[21]	train-rmse:130.309553
[22]	train-rmse:130.309542
[23]	train-rmse:130.309536
[24]	train-rmse:130.309531
[25]	train-rmse:130.309532
[26]	train-rmse:130.309532
[27]	train-rmse:130.309530
[28]	train-rmse:130.309530
[29]	train-rmse:130.309531
[30]	train-rmse:130.309528
[31]	train-rmse:130.309530
[32]	train-rmse:130.309530
[33]	train-rmse:130.309529
[34]	train-rmse:130.309530
[35]	train-rmse:130.309532
[36]	train-rmse:130.309529
[37]	train-rmse:130.309529
[38]	train-rmse:130.309530
[39]	train-rmse:130.309529
[40]	train-rmse:130.309530
[41]	train-rmse:130.309529
[42]	train-rmse:130.309533
[43]	train-rmse:130.309528
[44]	train-rmse:130.309530

[45]	train-rmse:130.309529
[46]	train-rmse:130.309530
[47]	train-rmse:130.309531
[48]	train-rmse:130.309531
[49]	train-rmse:130.309531
[50]	train-rmse:130.309531
[51]	train-rmse:130.309531
[52]	train-rmse:130.309531
[53]	train-rmse:130.309531
[54]	train-rmse:130.309531
[55]	train-rmse:130.309531
[56]	train-rmse:130.309531
[57]	train-rmse:130.309531
[58]	train-rmse:130.309531
[59]	train-rmse:130.309531
[60]	train-rmse:130.309531
[61]	train-rmse:130.309531
[62]	train-rmse:130.309531
[63]	train-rmse:130.309531
[64]	train-rmse:130.309531
[65]	train-rmse:130.309531
[66]	train-rmse:130.309531
[67]	train-rmse:130.309531
[68]	train-rmse:130.309531
[69]	train-rmse:130.309531
[70]	train-rmse:130.309531
[71]	train-rmse:130.309531
[72]	train-rmse:130.309531
[73]	train-rmse:130.309531
[74]	train-rmse:130.309532
[75]	train-rmse:130.309532
[76]	train-rmse:130.309532
[77]	train-rmse:130.309532
[78]	train-rmse:130.309532
[79]	train-rmse:130.309532
[80]	train-rmse:130.309532
[81]	train-rmse:130.309532
[82]	train-rmse:130.309532
[83]	train-rmse:130.309532
[84]	train-rmse:130.309532
[85]	train-rmse:130.309532
[86]	train-rmse:130.309532
[87]	train-rmse:130.309532
[88]	train-rmse:130.309532
[89]	train-rmse:130.309532
[90]	train-rmse:130.309532

[91]	train-rmse:130.309532
[92]	train-rmse:130.309532
[93]	train-rmse:130.309532
[94]	train-rmse:130.309532
[95]	train-rmse:130.309532
[96]	train-rmse:130.309532
[97]	train-rmse:130.309532
[98]	train-rmse:130.309532
[99]	train-rmse:130.309532
[100]	train-rmse:130.309532

Water Quality Index (WQI)

The Water Quality Index (WQI) measures water quality via an index value that indicates its overall appropriateness for diverse applications. The Water Quality Index (WQI) is an evaluative tool that measures the aggregate effect of specific water quality metrics on overall water quality. The average concentrations of ten physiochemical parameters—pH, electrical conductivity, total dissolved solids, dissolved oxygen, biochemical oxygen demand, iron, total alkalinity, total chloride, total hardness, and sulfate were employed to

Calculate the water Quality Index (WQI).

The Water Quality Index (WQI) values during the dry season were 50.10 for the Benin Formation (BF), 24.98 for the Ogwashi Asaba Formation (OAF), 20.18 for the Nsukka Formation (NF), 35.79 for the Alluvium Formation (AF), 79.77 for the Imo Clay Shale Formation (ICSF), and 55.94 for the False Bedded Sandstones Formation (FBSF), as presented in Tables 4.5 to 4.10. During the wet season, the WQI values recorded for the same formations were 35.04, 73.31, 27.54, 30.37, 86.99, and 108.95, as detailed in Tables 4.11 to 4.16.

Water samples collected during the dry season from the Ogwashi-Asaba and Nsukka Formations demonstrate enhanced water quality, as shown in Tables 4.6 and 4.7. In contrast, water samples from the Benin Formation and Alluvium Formation exhibit acceptable water quality, as seen in Tables 4.5 and 4.8, respectively. Table 4.10 demonstrates that samples from False Bedded Sandstone Formations display inferior water quality. The water sample from the Imo Clay shale formation has exceedingly low water quality during the dry season, as seen in Table 4.9. In contrast, during the rainy season, water samples from the Benin formation, Nsukka formation, and Alluvium formation exhibit high water quality, as shown in Tables 4.11, 4.13, and 4.14, whereas samples from the Ogwashi-Asaba formation reveal low water quality. The Water Quality Index obtained from groundwater samples in the Imo Clay Shale Formation, as shown in Table 4.15, reveals that the water quality in that area is extremely poor, and the water samples from the False Bedded Sandstones formation are considered unsuitable for consumption. This sector necessitates an innovative institutional economic strategy to tackle its current and future problems. The problems can be attributed to main pollutants and other deleterious elements that undermine water potability.

Table 4.4 Classification of water quality index (WQI) of drinking water

Water quality index level	Water quality status	Grading
0-25	Excellent water quality	A
26-50	good water quality	B
51-75	poor water quality	C
76-100	Very poor water quality	D
> 100	Unsuitable for drinking	E

Source: Ketata – Rokban et al. 2011.

Table 4.5: Calculation of WQI values for groundwater samples in Benin Formation during dry season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	5.63	8.5	0.1176	66.2353	7.7893
2.	Electrical Conductivity (EC)	215.00	750	0.00133	28.6667	0.03813
3.	Total Dissolved Solid (TDS)	139.75	1000	0.001	13.9750	0.01398
4.	Dissolved Oxygen (DO)	11.75	5	0.2	235	47
5.	Biochemical Oxygen Demand, (BOD)	6.95	5	0.2	139	27.8
6.	Iron(Fe) (mg/L)	0.10	0.3	3.33	33.3	111.10
7.	Total Alkalinity, ($CaCO_3$,)	5.00	200	0.005	2.5	0.0125
8.	Total Chloride, (Cl)	49.98	250	0.004	19.992	0.0799
9.	Total Hardness (TH)	77.70	500	0.002	15.54	0.0311
10.	Sulphate, (SO_4^{-2})	11.58	250	0.004	4.632	0.0185
				∑ Wi = 3.8649		∑(Qi x Wi) = 193.8694

WQI = ∑ (Q_i .Wi)/∑W_i = 205.059/3.8649 = 50.103

Table 4.6: Calculation of WQI values for groundwater samples in Ogwashi Asaba Formation during dry season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	5.70	8.5	0.1176	67.0588	7.8861
2.	Electrical Conductivity (EC)	91.00	750	0.0013	12.1333	0.0161
3.	Total Dissolved Solid (TDS)	66.50	1000	0.001	6.65	0.0067
4.	Dissolved Oxygen (DO)	7.58	5	0.2	151.6	30.32
5.	Biochemical Oxygen Demand, (BOD)	3.45	5	0.2	69	13.8
6.	Iron (Fe)	0.04	0.3	3.33	13.33	44.3889
7.	Total Alkalinity, ($CaCO_3$,)	19.65	200	0.005	9.825	0.0491
8.	Total Chloride, (Cl)	17.45	250	0.004	6.98	0.0279
9.	Total Hardness (TH)	109.25	500	0.002	21.85	0.0437
10.	Sulphate(SO_4^{-2})	5.00	250	0.004	2.00	0.008
				∑ Wi = 3.8649		∑(Qi x Wi) = 96.5465

WQI = ∑(Q_i.W_i)/∑W_i = 96, 5465/3.8649 = 24.9801

Table 4.7: Calculation of WQI values for groundwater samples in Nsukka Formation during dry season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	Ph	6.51	8.5	0.1176	76.5882	9.0068
2.	Electrical Conductivity (EC)	300.00	750	0.00133	40	0.0532
3.	Total Dissolved Solid (TDS)	400.00	1000	0.001	40	0.04
4.	Dissolved Oxygen (DO)	8.40	5	0.2	168	33.6
5.	Biochemical Oxygen Demand, (BOD)	6.00	5	0.2	120	24
6.	Iron, (Fe)	0.01	0.3	3.33	3.33	11.11

7.	Total Alkalinity, ($CaCO_3$,)	32.00	200	0.005	16	0.08
8.	Total Chloride, (Cl)	27.61	250	0.004	11.044	0.0442
9.	Total Hardness (TH)	160.18	500	0.002	32.036	0.0641
10.	Sulphate, (SO_4^{-2},)	8.50	250	0.004	3.40	0.0136
				∑ Wi = 3.8649		∑(Qi x Wi) = 78.0119

WQI = ∑ (Q_i. Wi)/∑W_i = 78.0119/3.8649 = 20.1847

Table 4.8: Calculation of WQI values for groundwater samples in Alluvium Formation during dry season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	5.78	8.5	0.1176	68	7.9968
2.	Electrical Conductivity (EC)	166.00	750	0.00133	22.133	0.0294
3.	Total Dissolved Solid (TDS)	107.90	1000	0.001	10.79	0.0108
4.	Dissolved Oxygen (DO)	7.20	5	0.2	144	28.8
5.	Biochemical Oxygen Demand, (BOD)	5.20	5	0.2	104	20.8
6.	Iron, (Fe)	0.072	0.3	3.33	24	79.920
7.	Total Alkalinity, ($CaCO_3$,)	35.00	200	0.005	17.5	0.0875
8.	Total Chloride, (Cl)	33.99	250	0.004	13.596	0.0544
9.	Total Hardness (TH)	152.81	500	0.002	30.562	0.0611
10.	Sulphate, (SO_4^{-2},)	6.71	250	0.004	2.684	0.0107
				∑ Wi = 3.8649		∑(Qi x Wi) = 138.3206

WQI = ∑ (Q_i .Wi)/∑W_i = 138.3206/3.8649 = 35.7889

Table 4.9: Calculation of WQI values for groundwater samples in Imo Clay Shale Formation during dry season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	5.31	8.5	0.1176	62.4705	7.3465
2.	Electrical Conductivity (EC)	210.20	750	0.00133	28.0267	0.0373
3.	Total Dissolved Solid (TDS)	234.13	1000	0.001	2.413	0.0024
4.	Dissolved Oxygen (DO)	10.60	5	0.2	212	42.40
5.	Biochemical Oxygen Demand, (BOD)	6.30	5	0.2	126	25.20
6.	Iron, (Fe)	0.21	0.3	3.33	70	233.10
7.	Total Alkalinity, ($CaCO_3$,)	45.80	200	0.005	22.90	0.1145
8.	Total Chloride, (Cl)	38.41	250	0.004	15.364	0.0615
9.	Total Hardness (TH)	130.30	500	0.002	26.06	0.0521
10.	Sulphate, (SO_4^{-2},)	3.68	250	0.004	1.472	0.0059
				∑ Wi = 3.8649		∑(Qi x Wi) = 308.3202

WQI = ∑ (Q_i. Wi)/∑W_i = 308.3202/3.8649 = 79.7744

Table 4.10: Calculation of WQI values for groundwater samples in False Bedded Sandstones Formation during dry season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	6.01	8.5	0.1176	70.706	8.3150
2.	Electrical Conductivity (EC)	156.70	750	0.00133	20.893	0.0278
3.	Total Dissolved Solid (TDS)	286.00	1000	0.001	28.60	0.0286
4.	Dissolved Oxygen (DO)	8.25	5	0.2	165	33.00
5.	Biochemical Oxygen Demand, (BOD)	4.82	5	0.2	96.4	19.28
6.	Iron (Fe)	0.14	0.3	3.33	15.00	155.411
7.	Total Alkalinity, ($CaCO_3$,)	18.36	200	0.005	9.18	0.0459
8.	Total Chloride, (Cl)	46.27	250	0.004	18.508	0.0740
9.	Total Hardness (TH)	180.93	500	0.002	36.186	0.0724
10.	Sulphate, (SO_4^{-2},)	10.25	250	0.004	4.1	0.0164
				∑ Wi = 3.8649		∑(Qi x Wi) = 216.1987

WQI = ∑ (Q_i.Wi)/∑W_i = 216.1987/3.8649 = 55.9390

Table 4.11: Calculation of WQI values for groundwater samples in Benin Formation during rainy season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	4.21	8.5	0.1176	66.2353	7.7893
2.	Electrical Conductivity (EC)	134.30	750	0.00133	49.5294	0.0659
3.	Total Dissolved Solid (TDS)	128.30	1000	0.001	12.80	0.0128
4.	Dissolved Oxygen (DO)	10.93	5	0.2	218.60	43.72
5.	Biochemical Oxygen Demand, (BOD)	4.28	5	0.2	85.60	17.12
6.	Iron(Fe) (mg/L)	0.06	0.3	3.33	20	66.60
7.	Total Alkalinity, ($CaCO_3$,)	4.10	200	0.005	2.05	0.0125
8.	Total Chloride, (Cl)	45.24	250	0.004	18.096	0.0724
9.	Total Hardness (TH)	72.50	500	0.002	14.50	0.0290
10.	Sulphate, (SO_4^{-2},)	10.28	250	0.004	4.112	0.0165
				∑ Wi = 3.8649		∑(Qi x Wi) = 135.4384

WQI = ∑ (Q_i .Wi)/∑W_i = 135.4384/3.8649 = 35.0432

Table 4.12: Calculation of WQI values for groundwater samples in Ogwashi Asaba Formation during rainy season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	5.34	8.5	0.1176	62.8235	7.3881
2.	Electrical Conductivity (EC)	60.10	750	0.00133	8.013	0.1068
3.	Total Dissolved Solid (TDS)	61.40	1000	0.001	6.14	0.0061
4.	Dissolved Oxygen (DO)	7.30	5	0.2	146.00	29.20
5.	Biochemical Oxygen Demand, (BOD)	3.35	5	0.2	67.00	13.40
6.	Iron,(Fe)	0.21	0.3	3.33	70.00	233.10

7.	Total Alkalinity, ($CaCO_3$,)	19.65	200	0.005	9.825	0.0491
8.	Total Chloride, (Cl)	17.82	250	0.004	7.128	0.0285
9.	Total Hardness (TH)	102.25	500	0.002	20.45	0.0409
10.	Sulphate(SO_4^{-2})	4.80	250	0.004	1.92	0.0077
				∑ Wi = 3.8649		∑(Qi x Wi) = 283.331

WQI = ∑(Q_i. Wi)/∑W_i = 383.331/3.8649 = 73.3077

Table 4.13: Calculation of WQI values for groundwater samples in Nsukka Formation during rainy season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight ($W_i=1/S_i$)	Quality Rating (Q_i= $100V_i/S_i$)	Q_i x W_i
1.	Ph	5.12	8.5	0.1176	60.24	7.0837
2.	Electrical Conductivity (EC)	260.00	750	0.00133	34.67	0.0462
3.	Total Dissolved Solid (TDS)	356.00	1000	0.001	35.6	0.0356
4.	Dissolved Oxygen (DO)	7.80	5	0.2	156	31.2000
5.	Biochemical Oxygen Demand, (BOD)	5.60	5	0.2	112	22.4000
6.	Iron, (Fe)	0.041	0.3	3.33	13.67	45.5100
7.	Total Alkalinity, ($CaCO_3$,)	31.00	200	0.005	15.5	0.0775
8.	Total Chloride, (Cl)	21.42	250	0.004	8.568	0.0343
9.	Total Hardness (TH)	130.18	500	0.002	26.036	0.0521
10.	Sulphate, (SO_4^{-2},)	7.78	250	0.004	3.112	0.0125
				∑ Wi = 3.8649		∑(Qi x Wi) = 106.4519

WQI = ∑ (Q_i. Wi)/∑W_i = 106.4519/3.8649 = 27.5433

Table 4.14: Calculation of WQI values for groundwater samples in Alluvium Formation during rainy season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight ($W_i=1/S_i$)	Quality Rating (Q_i= $100V_i/S_i$)	Q_i x W_i
1.	pH	5.20	8.5	0.1176	61.1765	7.1944
2.	Electrical Conductivity (EC)	130.00	750	0.00133	17.333	0.0231
3.	Total Dissolved Solid (TDS)	101.20	1000	0.001	10.12	0.0101
4.	Dissolved Oxygen (DO)	6.41	5	0.2	128.2	25.6400
5.	Biochemical Oxygen Demand, (BOD)	4.34	5	0.2	86.8	17.3600
6.	Iron, (Fe)	0.67	0.3	3.33	20.1	66.933
7.	Total Alkalinity, ($CaCO_3$,)	33.50	200	0.005	16.75	0.0838
8.	Total Chloride, (Cl)	33.99	250	0.004	13.596	0.0544
9.	Total Hardness (TH)	132.07	500	0.002	26.414	0.0528
10.	Sulphate, (SO_4^{-2},)	6.34	250	0.004	2.536	0.0101
				∑ Wi = 3.8649		∑(Qi x Wi) = 117.3617

WQI = ∑ (Q_i .Wi)/∑W_i = 117.3617/3.8649 = 30.3660

Table 4.15: Calculation of WQI values for groundwater samples in Imo Clay Shale Formation during rainy season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	4.75	8.5	0.1176	55.8824	6.5718
2.	Electrical Conductivity (EC)	185.80	750	0.0013	24.773	0.0322
3.	Total Dissolved Solid (TDS)	230.45	1000	0.001	23.045	0.0231
4.	Dissolved Oxygen (DO)	10.45	5	0.2	522.5	104.5000
5.	Biochemical Oxygen Demand, (BOD)	5.15	5	0.2	103	20.6000
6.	Iron, (Fe)	0.18	0.3	3.33	60	199.800
7.	Total Alkalinity, ($CaCO_3$,)	42.60	200	0.005	21.3	0.1065
8.	Total Chloride, (Cl)	35.20	250	0.004	14.08	0.0563
9.	Total Hardness (TH)	124.60	500	0.002	24.92	0.0498
10.	Sulphate, (SO_4^{-2})	3.56	250	0.004	1.424	0.0057
				∑ Wi = 3.8649		∑(Qi x Wi) = 336.1891

WQI = ∑ (Q_i. Wi)/∑W_i = 336.1891/3.8649 = 86.9852

Table 4.16: Calculation of WQI values for groundwater samples in False Bedded Sandstones Formation during rainy season

S/N	Parameter	Mean Monitored Value (V_i)	WHO Maximum Standard (S_i)	Unit weight (W_i=1/S_i)	Quality Rating (Q_i= 100V_i/S_i)	Q_i x W_i
1.	pH	5.62	8.5	0.1176	66.118	7.7754
2.	Electrical Conductivity (EC)	130.40	750	0.00133	17.387	0.0226
3.	Total Dissolved Solid (TDS)	246.00	1000	0.001	24.60	0.0246
4.	Dissolved Oxygen (DO)	8.02	5	0.2	160.4	32.0800
5.	Biochemical Oxygen Demand, (BOD)	3.75	5	0.2	75.0	15.0000
6.	Iron, (Fe)	0.33	0.3	3.33	110.0	366.300
7.	Total Alkalinity, ($CaCO_3$,)	16.32	200	0.005	9.18	0.0459
8.	Total Chloride, (Cl)	32.55	250	0.004	13.02	0.0521
9.	Total Hardness (TH)	175.68	500	0.002	35.136	0.0727
10.	Sulphate, (SO_4^{-2})	10.12	250	0.004	4.048	0.0169
				∑ Wi = 3.649		∑(Qi x Wi) = 421.0902

WQI = ∑ (Q_i. Wi)/∑W_i = 421.0902/3.8649 = 108.9524

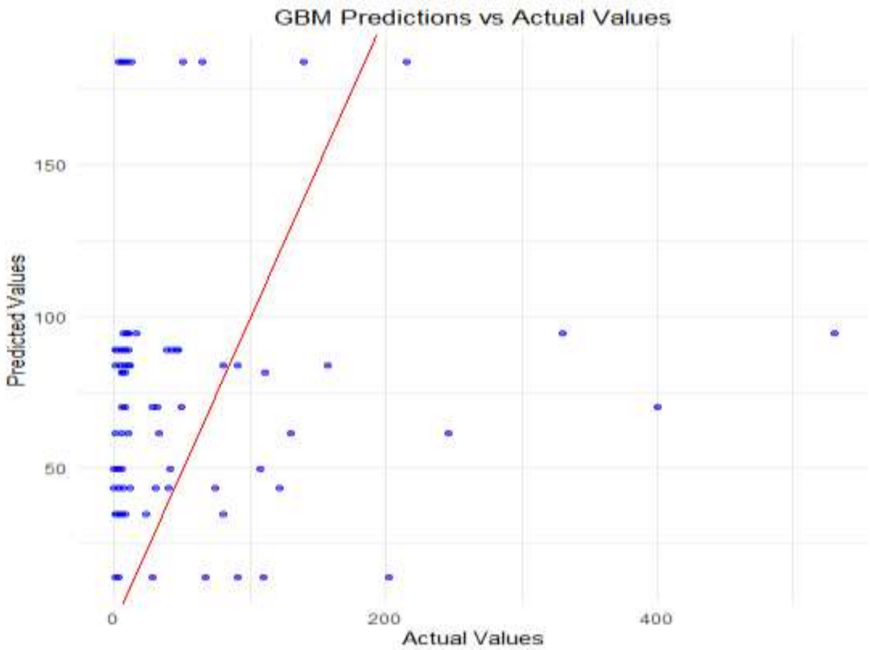

Figure 4.1. Graph of Actual and Predicted using the XGBoost Model.

V. Discussion

The quality of groundwater is a crucial determinant of human health, especially in countries reliant on natural resources, such as Nigeria. Groundwater is a vital and precious natural resource, anticipated to be devoid of contaminants. This water supply is frequently polluted by numerous contaminants originating from agricultural, industrial, and domestic sources. The fast increase in population and industry necessitates an examination of groundwater quality due to its susceptibility to municipal and industrial waste disposal.This study examined the spatiotemporal variability of groundwater quality across six geological zones in Imo State. It assessed the physiochemical characteristics of groundwater samples from the Benin Formation (BF), Ogwashi Asaba Formation (OAF), Nsukka Formation (NF), Alluvium Formation (AF), Imo Clay Shale Formation (ICSF), and False Bedded Sandstone Formation (FBSF). The mean concentrations of total dissolved solids, chloride, nitrate, sulfate, total hardness, and electrical conductivity were heightened throughout the dry season relative to the rainy season, although the mean concentrations of potassium and bicarbonate were higher in the wet season. Based on the outcomes of this study, the subsequent recommendations are put forth.

i. Periodic monitoring and remediation of acidic and nitrate-rich water are advised to prevent future contamination and ensure the sustainable utilization of groundwater resources.

ii. Additional research may be conducted regarding other significant climatic variables, including soil and air temperature, as well as solar radiation, which could affect aquifer conditions and dictate the depletion and degradation of groundwater.

iii. Collaborative efforts among the state environmental protection agency, the water resources ministry, the sanitation agency, and waste management organizations is crucial for developing and implementing a framework that protects water resources, enhances community access to potable water, and guarantees sustainable waste management.

VI. Conclusions

The findings indicate that during the dry season, groundwater samples from Ogwashi Asaba and Nsukka formations exhibit excellent water quality, while samples from the Benin and Alluvium formations demonstrate good water quality. Conversely, samples from the False Bedded Sandstones and Imo Clay Shale formations are characterized by poor water quality according to national and international indices and standards. This indicates that water from these two places necessitates treatment beforehand. The results from the rainy season showed that water samples from the Benin formation, Nsukka formation, and Alluvium formation exhibited high water quality, whereas samples from the Ogwashi Asaba and Imo Clay Shale formations demonstrated poor and extremely bad water quality, respectively. The sample from the false Bedded Sandstone formation is unfit for drinking purposes. This pertains to identifiable, indiscriminate releases of industrial wastewater and sewage. Conclusively, the groundwater quality across the six geological zones of Imo State was evaluated by analyzing various water quality parameters using laboratory techniques. Secondly, the results obtained were juxtaposed with the permitted limits established by WHO, BIS, and FMEnv.During the dry season, it was found that dissolved oxygen levels above the permissible limit of 7.5 in all formations, except for the Alluvium formation, which recorded exactly 7.5. Chemical Oxygen Demand exceeded the WHO allowed limit in all formations; Phosphate levels beyond the allowable limit in the Benin Formation, Nsukka Formation and Alluvium Formation. During the rainy season, Dissolved Oxygen was higher in all except Alluvium formation, COD exceeded the allowable limit in all the formations except Ogwashi Asaba formation; potassium was higher than the limit in Alluvium formation and phosphate was higher than the allowable limit except in Ogwashi Asaba Formation and False Bedded Sandstone Formations. Thirdly, the data collected demonstrated that TDS, Chlorides, Nitrates, Sulphate, Total Hardness and Electrical Conductivity increased in dry season whereas Potassium and Bicarbonate were higher in wet season. Fourthly, the Water Quality Index (WQI) of the tested water samples was calculated, and they yielded the following values: The Water Quality Index (WQI) for the dry season was 50.10, 24.98, 20.18, 35.79, 79.77, and 55.94 for BF, OAF, NF, AF, ICSF, and FBSF, respectively. In contrast, the WQI for the wet season was 35.04, 73.30, 27.54, 30.37, 86.98, and 108.95 for BF. OAF, NF, AF, ICSF, and FBSF respectively. Finally, the XGBOOST model was employed to analyze the variances. This was trained using a 70-30 ratio where 70% was for calibration (training) and 30% for validation (testing), this resulted in a Root Mean Square Error (RMSE) value of 142.8292 that later decreased to 130.3095 at the final iteration after having undergone hundred (100) iterations. The decreased value of RMSE from 142.8292 to 130.3095 indicates convergence and limited potential for improvement with further iterations. Invariably, at this point of final iteration that was optimized, the system can be predicted.

CONTRIBUTIONS TO KNOWLEDGE

This work enhances knowledge by improving the understanding of the hydrogeology of geological zones. The produced spatial variability will aid water resource managers and policymakers in formulating guidelines to combat future pollution and in the wise control of groundwater resources for both agricultural and consumable applications in the research areas.

iii. Forecasts of groundwater quality and quantity will assist in pinpointing suitable agricultural regions and preventing over-extraction of water, when salt levels relative to
Calcium and magnesium concentrations could significantly rise.

References

[1]. Vasanthavigar M., Srinivasamoorthy K., Vijayaragavan K., Rajiv Ganthi, R., Chidambaram, S., Anandhan, P., Manivannan, R. and Vasudevan, S. (2010). Application of water quality index for groundwater quality assessment: Thirumanimuttar sub-basin, Tamilnadu, India. Environmental Monitoring and Assessment, 171(1–4), pp. 595–609.

[2]. Vutukuru, S. (2005) Acute Effects of Hexavalent Chromium on Survival, Oxygen Consumption, Hematological Parameters and Some Biochemical Profiles of the Indian Major Carp, Labeo rohita. International Journal of Environmental Research and Public Health, 2, 456-462. https://doi.org/10.3390/ijerph2005030010

[3]. Wang, Y.B., Liu, C.W., Liao, P.Y., & Lee, J.J. (2014). Spatial pattern assessment of river water quality: implications of reducing the number of monitoring stations and chemical parameters. Environmental monitoring and assessment, 186(3), 1781-1792. https://doi.org/10.1007/s10661-013-3492-9

[4]. "Where is Earth's Water?" (2020) www.usgs.gov.

[5]. World Health Organization (1997). "Guidelines for drinking-water quality". Surveillance and Control of Community Supplies.

[6]. Zhang Q., Wang L., Wang H., Zhu X., Wang L. (2020). Spatio-temporal Variation of Groundwater Quality and Source Apportionment using Multivariate Statistical Techniques for the Hutuo RiverAlluvial-Pluvial Fan, China. International Journal of Environmental Research and Public Health,17, 1055; doi:10.3390/ijerph17031055.

[7]. Zhang, Q.; Miao, L.; Wang, H.; Hou, J.; Li, Y. (2020). How Rapid Urbanization Drives Deteriorating Groundwater Quality in a Provincial Capital of China. Pol. J. Environ. Stud., 29, 441–450.

[8]. Zhang, Q.; Wang, H.; Wang, Y.; Yang, M.; Zhu, L. (2017). Groundwater quality assessment and pollution source apportionment in an intensely exploited region of northern China. Environ. Sci. Pollut. Res., 24, 16639–16650.

[9]. Zheng, L.Y., Yu, H.B., & Wang, Q.S. (2016). Application of multivariate statistical techniques in assessment of surface water quality in Second Songhua River basin, China. Journal of Central South University, 23, 1040-1051. https://doi.org/10.1016/j.jher.2012.10.003