ASSESSMENT OF GROUNDWATER POTENTIAL USING MACHINE LEARNING TECHNIQUES: A CASE STUDY OF THE EASTERN PART OF AHMADU BELLO UNIVERSITY, ZARIA, NORTHWEST NIGERIA

NOUHOU I., UMEGO M. N.[1], AJAYI C. O.[2] and SULE P. O.[2]
Université Dan Dicko Dankoulodo (UDDM), MARADI, Republique du Niger, Département de Physique, Faculté des Sciences et Techniques (FST)
[1] Nnamdi Azikwe University (UNIZIK), AWKA, Federal Republic of Nigeria
Department of Physics, Faculty of Science
[2]Ahmadu Bello University (ABU), ZARIA, Federal Republic of Nigeria
Department of Physics, Faculty of Science
Corresponding Author Email: nouhou.ibrahim@uddm.edu.ne
nuhu.brhm@gmail.com

Abstract
This study evaluates groundwater potential in the eastern part of Ahmadu Bello University, Zaria, using machine learning techniques integrated with geoelectric data. Thirty-nine Vertical Electrical Sounding (VES) points were analysed using the Schlumberger configuration to determine aquifer resistivity and thickness. Optimal groundwater potential was identified in zones with resistivity values between 50–300 Ω-m and aquifer thickness ranging from 20–40 m, corresponding to weathered layers conducive to water storage. K-Means clustering, supported by the elbow method, categorized groundwater potential into high, medium, and low levels, highlighting significant spatial variations. High-potential zones, such as VES 8 and 19, demonstrated substantial aquifer thickness and moderate resistivity, while low-potential zones, including VES 31 and 39, showed minimal thickness and unfavourable resistivity conditions. Violin and contour plots revealed strong correlations between resistivity, thickness, and water yield potential, supporting targeted exploration strategies. The study recommends prioritizing high-potential zones for sustainable groundwater development while incorporating advanced machine learning models and further geophysical surveys to enhance prediction accuracy. These findings underscore the utility of machine learning in optimizing groundwater resource management in geologically similar terrains.

Keywords: Groundwater potential, Machine Learning, K-Means Clustering, Resistivity and Thickness.

Introduction
Groundwater is a crucial source of drinking water and also a viable source for meeting domestic and industrial water needs (Ifeany *et al.*, 2020). The yield of boreholes is often low and short-term water supply sustainability in Zaria environs (Sule & Inichinbia, 2018). This challenge can be attributed to inadequate groundwater investigation before drilling of boreholes. So, understanding the aquifer parameters which include resistivity and thickness is essential for determining groundwater potential (Onowola *et al.*, 2020).

Electrical resistivity method can be used to investigate the subsurface resistivity distribution by passing current into the subsurface via a pair of electrodes known as current electrodes and recording the potential difference via another pair of electrodes known as potential electrodes. However, the data obtained from the investigation carried out with the use of this method will be further processed and interpreted to estimate geoelectric parameters which include resistivity and thickness of aquifer layers. Thirty-nine (39) vertical electrical sounding (VES) points were established with the use of Schlumberger configuration during data acquisition for this research in the study area.

Machine learning (ML) has emerged as a potent tool in several discipline in which geoscience is not exempted and the employment of data analysis techniques, facilitated by the application of machine learning (ML) has the potential to drive significant breakthroughs across industries and academic fields (David and Bery, 2024). The clustering analysis utilised in this K-Means clustering which is known for its effectiveness in identifying cluster centroids (Bholowalia & Kumar, 2014). In this study, the elbow method is used to determine the optimal number of clusters.

The elbow method assesses the variance explained relative to the number of clusters. The core idea is that the optimal number of clusters should be identified at the point where adding another cluster does not significantly improve the model's performance. A visual representation illustrates the relationship between the number of clusters and the variance explained. Initially, clusters provide substantial information, but there comes a point where the additional benefit diminishes, resulting in a noticeable change in the graph's slope. Thus, the optimal number of clusters, denoted as, k, is determined using the elbow method.

Location and Geology of the Study Area

The study site is located within the Ahmadu Bello University main campus in Zaria. It is bounded by latitude 11°091 N and 11°101 N longitude 7°381 E and 7°391 E. It forms part of the plain that extends from Lake Chad to Sokoto and Northward from Southern Kaduna into the Republic of Niger (Ifeanyi *et al.,* 2020). The two distinct seasons in the Zaria are the wet season which lasts from April to October and dry season lasts from November to March. Zaria is underlain by Precambrian rocks typical of Nigeria basement complex in which the north-eastern part where this study forms a part is underlain by biotite granite-gneiss. The geology of Zaria is shown in Figure 1, (Sule and Inichinbia, 2018).

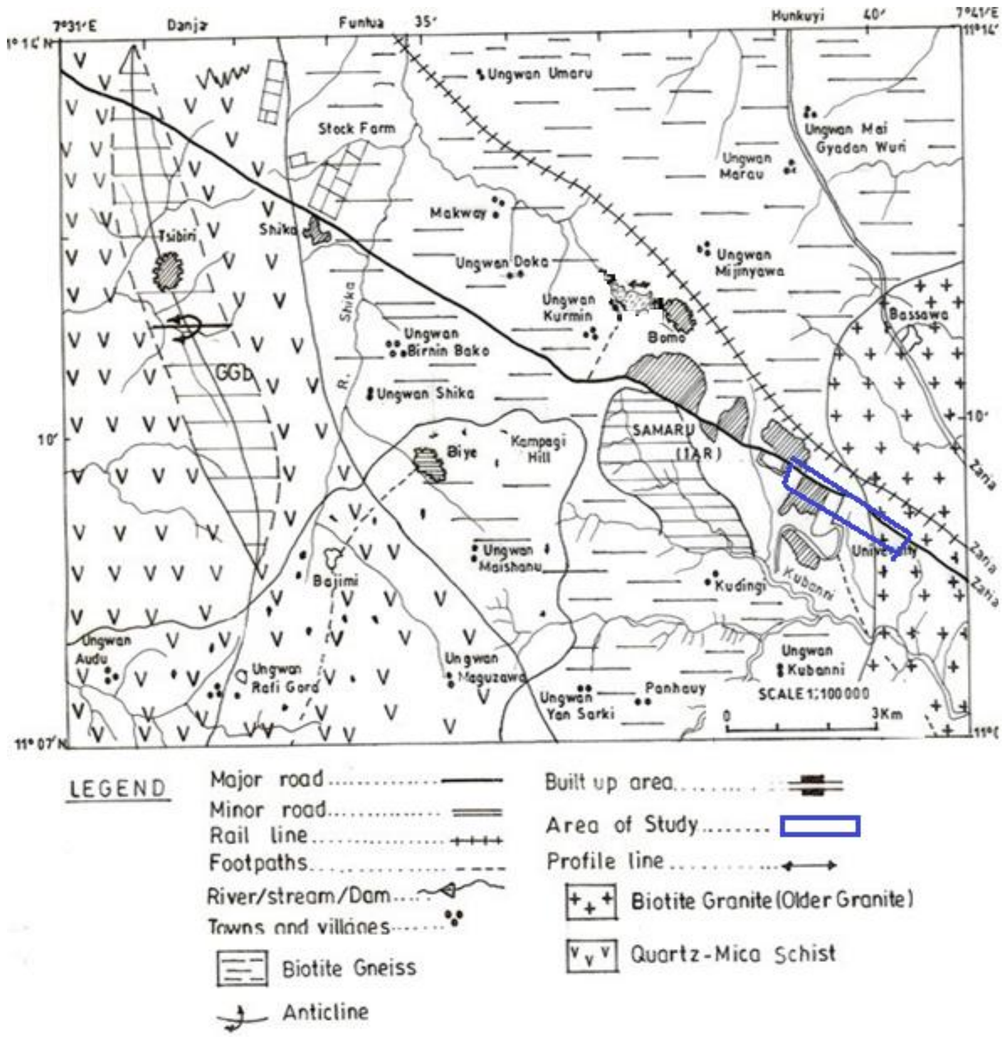

Figure 1: Geology Map of Zaria (After Sule and Inichinbia, 2018)

Aquifer Potential as a Function of Resistivity

Aquifer potential as a function of resistivity is presented in Table 1. It can be deduced from the table that maximum groundwater potential can be obtained from an aquifer layer of resistivity values which are within the range of 50 Ω-m and 300 Ω-m. This can be attributed to complete weathering which result in the formation of sand. Limited groundwater can be obtained in layers of resistivity values which are within the range of 300 Ω-m and 400 Ω-m and this can be attributed to partial weathering of the layers. However, negligible potential is found in layers with resistivity values greater than 400 Ω-m in the study area.

Table 1: Aquifer Potential as a Function of Resistivity of the Study Area (Modified after Fajana, 2020)

Resistivity (Ohm-m)	Aquifer Characteristics
<50	Clay (limited aquifer potential)
50-200	Optimum weathering and groundwater potential
200-300	Medium aquifer conditions and potential
300-400	Limited weathering and poor potential
>400	Negligible potential

Results and Discussions

The resistivity values and thickness of aquifer layers in the study area are presented in Table 2. It shows the range of resistivity and thickness values for each VES points and this values were further analysed with the use of Machine Learning techniques categorise groundwater potential in the study area into low, medium and high potential.

Table 2: Resistivity and Thickness of Aquifer Layers

VES Points	Resistivity (Ohm-metres)	Thickness (Metres)
1	175	24.8
2	135	11.2
3	449.5	30.18
4	147.1	29.7
5	160	22.74
6	354.8	27.62
7	451.2	25.61
8	176.1	33.74
9	111.3	25.61
10	132.7	11.68
11	138.6	34.96
12	48.67	6.25
13	250	22.8
14	223.1	38.67
15	49.96	14.17
16	157.6	23.32
17	66.38	23.02
18	125	29
19	100	27.98
20	17	8.12
21	39.61	5.66
22	27	5.31
23	81	33.8
24	45	4.09
25	222	26.5
26	146	14.98
27	78	11.33
28	87	13.67
29	91.2	23.06
30	72	9.94
31	43	1.36
32	150	5.9
33	98	4.6
34	94	20.4
35	106	6.31
36	88	6
37	68	4
38	38	4.4
39	145	3.63

Elbow Method of Clustering

The Elbow Plot, shown in Figure 2, illustrates the relationship between the number of clusters (k) and inertia, which is the sum of squared distances between each point and its cluster centroid. The x-axis represents the number of clusters, ranging from 1 to 10, while the y-axis shows inertia, demonstrating how well data points are grouped. Initially, the curve decreases sharply as more clusters are added, indicating a significant reduction in within-cluster variance; however, after reaching a certain point—known as the "elbow"—the rate of decrease slows considerably, suggesting that additional clusters do not significantly enhance clustering performance. In this case, the elbow appears at k = 3, indicating that three clusters are optimal for this dataset, which is especially relevant for attributes like resistivity and thickness. This approach helps prevent overfitting by limiting the number of clusters while still allowing for meaningful groupings within the data.

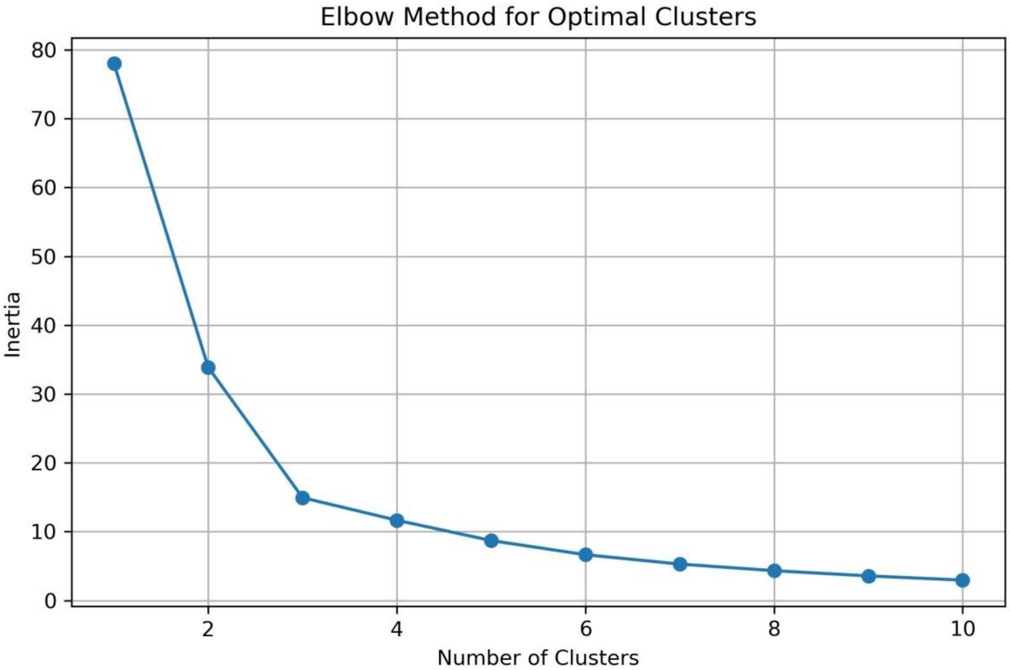

Figure 2: Elbow Clustering Plot

Resistivity and Thickness Distribution by Groundwater Potential

The violin plots illustrate the distributions of resistivity (Ωm) and thickness (m) based on groundwater potential levels (High, Medium, and Low). In the resistivity plot, areas with high groundwater potential exhibit a wide range of resistivity values, typically higher than those for medium and low groundwater potential. Low potential areas show the narrowest range, concentrated at lower resistivity values, indicating poor aquifer quality which may be due to high clay content. Similarly, the thickness distribution indicates that high groundwater potential zones are associated with thicker aquifer layers, generally around 30 meters. Medium potential zones have moderate thickness, while low potential areas have thinner layers concentrated below 20 meters. These patterns suggest a strong correlation between aquifer thickness, resistivity, and groundwater potential, which can guide groundwater exploration.

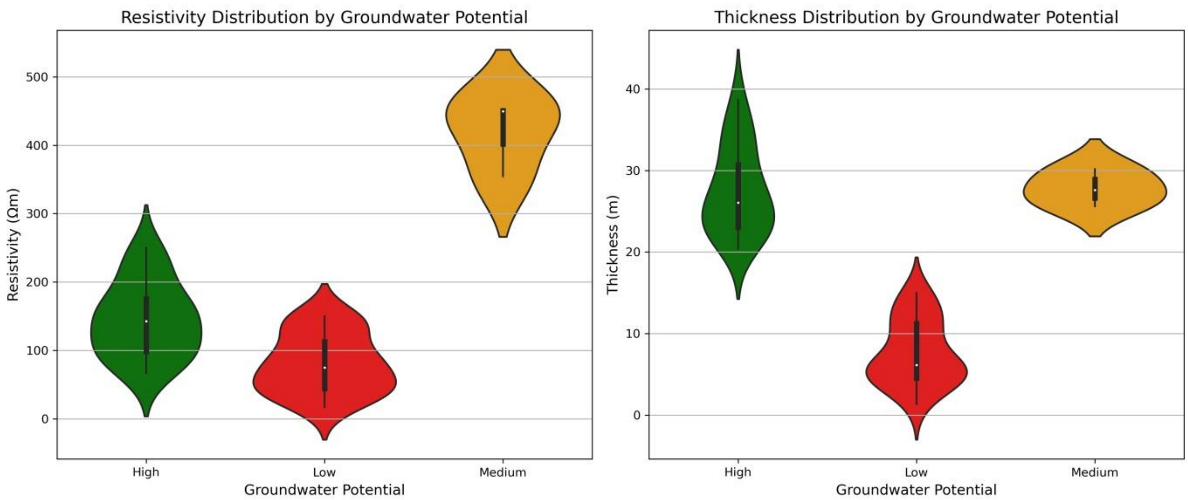

Figure 3: Violin Plots of Resistivity and Thickness Distribution by Groundwater Potential

Groundwater Potential Clustering: Figure 4 represents the clustering of groundwater potential (High, Medium, and Low) based on resistivity (Ωm) and thickness (m) of the aquifer. High groundwater potential (green dots) is characterized by higher aquifer layers resistivity values, ranging from approximately 100–300 Ωm (indicating weathered layers), and greater aquifer thickness, generally between 20–40 meters. Medium potential zones (orange squares) show moderate resistivity values, around 200–400 Ωm (indicating partially weathered/fractured layer), and aquifer thickness between 10–30 meters. Low potential areas (red crosses) are clustered at lower resistivity values of aquifer layers (which may be due to the availability of clay), approximately 0–100 Ωm, with aquifer thickness mostly below 15 meters. The clustering highlights distinct relationships between resistivity and thickness for different groundwater potential levels. These patterns can aid in identifying zones suitable for groundwater extraction based on geophysical parameters.

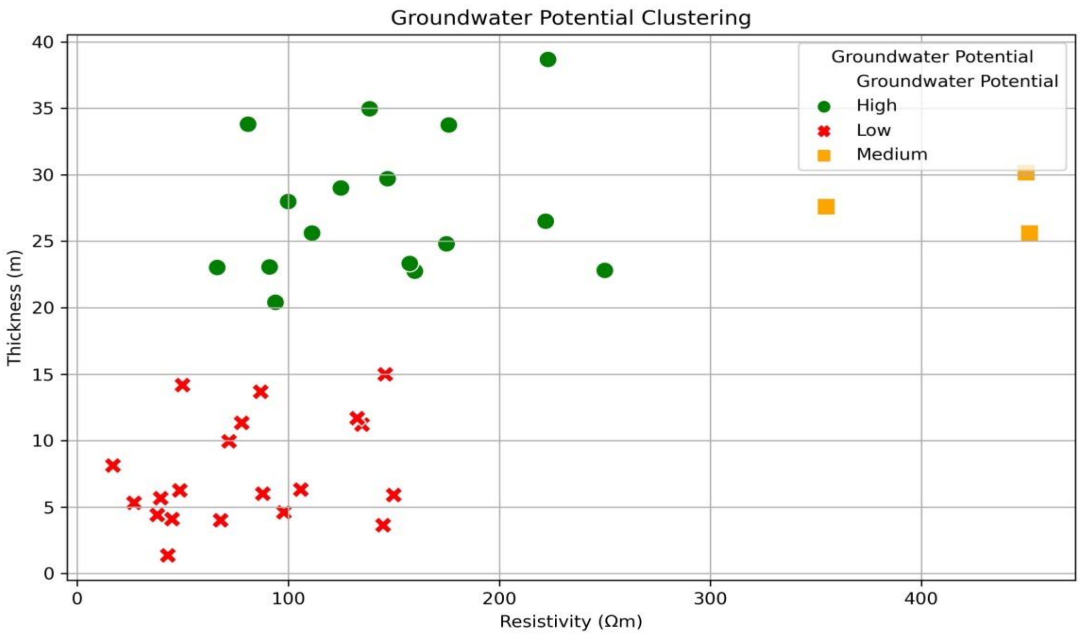

Figure 4: Groundwater Potential Clustering

Interactive Groundwater Potential Visualisation

The interactive plot visualizes groundwater potential (High, Medium, and Low) based on resistivity (Ω-m) and thickness (m) of the aquifer. High groundwater potential (blue dots) is associated with higher resistivity values of aquifer layers, ranging from 100–300 Ω-m, and thicker aquifer layers, mostly between 20–40 meters. Medium potential (green dots) shows moderate resistivity, ranging from 200–400 Ω-m, and aquifer thickness primarily between 10–30 meters. Low potential (red dots) clusters at lower resistivity values, typically below 100 Ω-m, and thinner aquifer layers under 15 meters. The distribution shows clear separation between groundwater potential levels, emphasizing that higher resistivity and thickness correlate with better groundwater potential. The interactive nature allows users to explore specific data points and better understand the spatial distribution of these parameters for groundwater evaluation.

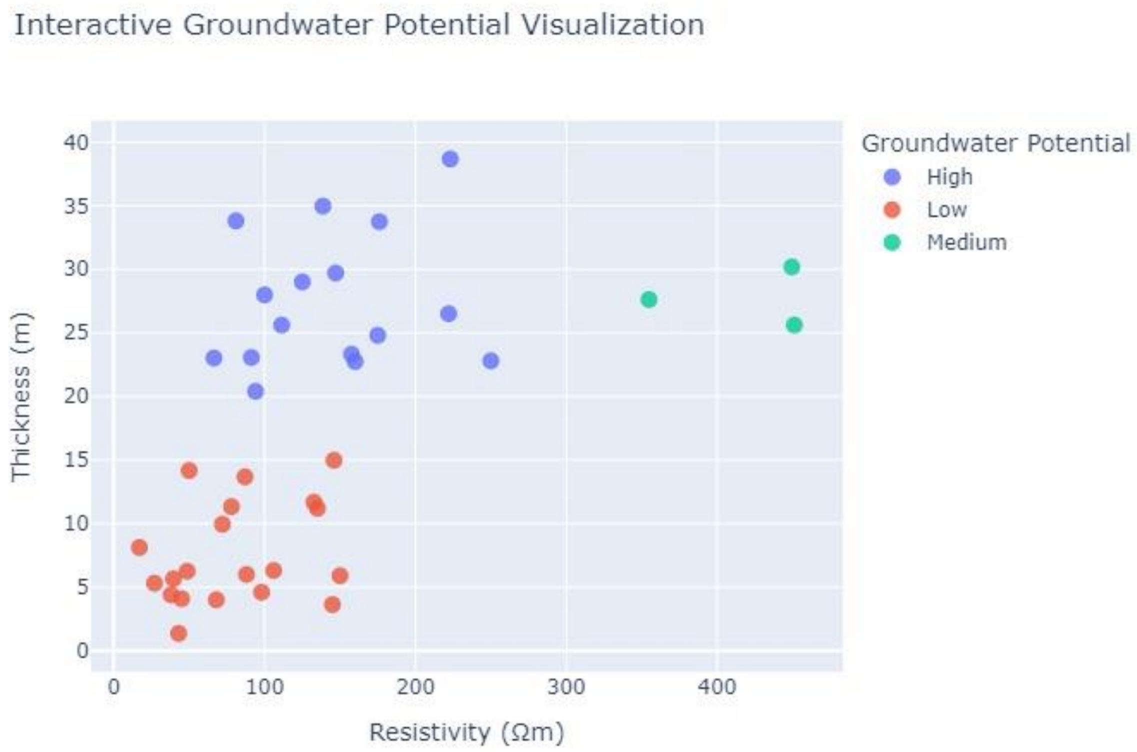

Figure 5: Interactive Groundwater Potential Visualisation

Contour Plot of Groundwater Potential

The contour plot, shown in Figure 6, illustrates groundwater potential based on resistivity and aquifer thickness derived from Vertical Electrical Sounding (VES) data. The X-axis represents resistivity (Ω-m), while the Y-axis indicates aquifer thickness (m), with colours showing groundwater potential clusters. Red regions indicate high groundwater potential, associated with moderate resistivity (50–150 Ω-m) and substantial aquifer thickness (20–35 m). Blue regions signify low potential, typically corresponding to low thickness (<10 m) or very high resistivity (>300 Ω-m). VES points are labelled and plotted, highlighting areas with varying water-bearing capacities. High-potential zones, like VES 8 and 19, are ideal for groundwater development, while low-potential zones, such as VES 31 and 39, are less suitable. Intermediate zones with moderate conditions serve as transitional areas between high and low potentials. These zones are represented by lighter shades of red or blue on the contour plot. These zones include VES points such as 9, 13, 16, 17, 25, and 28, located where resistivity ranges between 100–200 Ω-m and aquifer thickness is between 10–20 m. These areas may have moderate water-bearing properties. This analysis aids in identifying optimal locations for groundwater exploration and resource management.

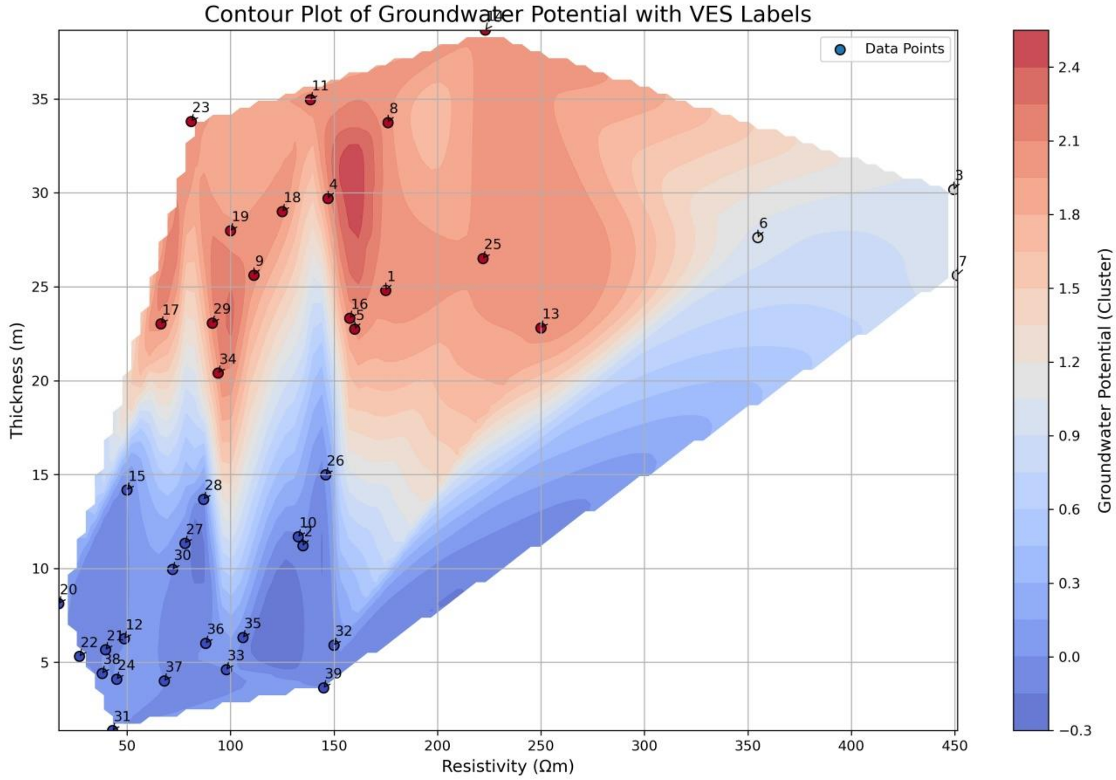

Figure 6: Contour Plot of Groundwater Potential with VES Points

Conclusion

This research concludes that groundwater potential in the eastern part of Ahmadu Bello University, Zaria, can be effectively assessed using machine learning techniques combined with geoelectric data. The study demonstrated that resistivity and aquifer thickness are critical parameters influencing groundwater potential, with optimal zones characterized by moderate resistivity (50–300 Ωm) and substantial thickness (20–40 m). The application of K-Means clustering and the elbow method successfully categorized groundwater potential into high, medium, and low levels, providing a clear correlation between these parameters and potential water yield. High groundwater potential areas correspond to weathered layers with significant aquifer thickness, while low-potential zones are associated with lower thickness and higher clay content. The findings emphasize the value of integrating machine learning in geophysical analysis to enhance resource management. Based on the results, it is recommended that future groundwater exploration in the study area prioritize high-potential zones, such as VES points 8 and 19, for sustainable water development. Additionally, incorporating advanced machine learning models and further geophysical surveys can improve the accuracy of groundwater potential predictions in similar terrains.

References

David D.M. & Bery A.A. (2024). Geophysical Data Optimization with Machine Learning for Lithological Characterization,1-12.

Fajana A.O. (2020). Groundwater Aquifer Potential using Electrical Resistivity Method and Porosity Calculation: *NRIAG Journal of Astronomy and Geophysics*, 9:1, 168-175, DOI: 10.1080/20909977.2020.1728955.

Ifeanyi O., Aduojo A.A., Olaide K.F., Stephen N.C. & John R.O. (2020). Evaluation of Groundwater Potential in Part of Ahmadu Bello University, Zaria, Northwest, Nigeria, 1-13.

Onawola B.R., Olatunji S., Ologe O. & Jimoh R.O. (2021). Determination of Aquifer Parameters from Resistivity Data: A Case of University of Ilorin Campus, Northcentral Nigeria: *Tanzania Journal of Science* 47(1): 91-103,

Sule P.O. & Inichinbia (2018) 2D Geo-electrical Imaging for Groundwater Exploration in Shika, along Zaria-Funtua Road North of Ahmadu Bello University Main Campus, Samaru. Zaria: *Journal of Applied Science and Environment Management*, 22(4):609–612.