RESEARCH ARTICLE

A Machine Learning Framework for the Investigation of Energy-Critical Mineralized Geologic Structures from Gravity and Magnetic Datasets: Implications for Sustainable Exploration

Ema Abraham[1,*], **George-Best Azuoko**[1], **Ayatu Usman**[1], **Ikechukwu Onwe**[1], **Iheanyi Ikeazota**[1], **Cyril Afuwai**[2], **Ene Obande**[3], **Stanley Elom**[1] and **Rock Onwe**[1]

[1] Department of Geology/Geophysics, Alex Ekwueme Federal University, Ndufu-Alike Ikwo, P.M.B. 1010, Abakaliki, Ebonyi, Nigeria
[2] Department of Physics, Kaduna State University, P.M.B. 2339, Kaduna, Nigeria
[3] Department of Quality Assurance, National University Commission, Abuja, Nigeria

Abstract

Mineral exploration faces challenges from complex geological architectures and subtle geophysical expressions of mineralization. This study proposes a joint gravity–magnetic machine learning framework integrating magnetic anomaly, analytic signal, gravity, and Source Parameter Imaging depth estimates to enhance mineralization prediction in Nigeria's Middle Benue Trough and adjoining basement terrain. Five supervised algorithms were evaluated, with Random Forest achieving the highest performance (accuracy=0.954, precision=0.889, recall=0.819, macro-F1=0.850, ROC-AUC=0.978). Correlation analysis revealed low feature redundancy, while unsupervised clustering confirmed structural partitions consistent with mapped fault systems. Ablation studies identified analytic signal as the most influential predictor; however, gravity and depth features contributed essential complementary information, increasing predictive accuracy by over 10% when combined with magnetic data. The resulting mineralized structure map aligns with conventional interpretations while delineating previously unrecognized targets in subdued magnetic response areas. This framework provides an objective, geologically defensible tool for energy-critical mineral targeting, demonstrating substantial improvements over traditional methods while minimizing environmental disturbance through enhanced exploration efficiency. The approach is directly applicable to sustainable resource development in similar terranes worldwide.

Keywords: energy-critical minerals, mineral resources, magnetic, gravity, machine learning.

Submitted: 30 January 2026
Accepted: 11 February 2026
Published: 05 March 2026

Vol. 2, **No.** 2, 2026.
10.62762/JGEE.2026.309205

***Corresponding author:**
✉ Ema Abraham
ema.abraham@funai.edu.ng

Citation
Abraham, E., Azuoko, G. B., Usman, A., Onwe, I., Ikeazota, I., Afuwai, C., Obande, E., Elom, S., & Onwe, R. (2026). A Machine Learning Framework for the Investigation of Energy-Critical Mineralized Geologic Structures from Gravity and Magnetic Datasets: Implications for Sustainable Exploration. *Journal of Geo-Energy and Environment*, 2(2), 95–117.

1 Introduction

Mineral exploration in Nigeria has gained significant momentum over the past two decades, driven by diverse geological framework and the increasing global demand for critical minerals. Northern Nigeria, characterized by extensive Precambrian basement complexes and sedimentary sequences, hosts numerous mineralized zones that remain underexplored despite their economic potential [1–6]. The region's geological complexity, spanning from the Nigerian Basement Complex to the Chad Basin (Figure 1), presents both opportunities and challenges for modern exploration methodologies.

The minerals targeted in this study, including rare earth elements (REEs), critical metals (tantalum, niobium, tin), and industrial minerals, are essential for the global energy transition and renewable energy technologies. These resources serve as critical components in solar photovoltaic panels, wind turbine magnets, lithium-ion battery electrodes, energy-efficient lighting systems, and advanced power electronics [7, 8]. The International Energy Agency estimates that demand for critical minerals could increase by 400-600% by 2040 under clean energy scenarios, with specific minerals like lithium, cobalt, and rare earths experiencing even higher growth rates [9]. Furthermore, coal deposits identified in the Lafia Formation, while representing traditional energy resources, require careful exploration strategies to balance energy security with environmental sustainability goals and climate commitments. The machine learning framework developed here contributes to environmentally responsible exploration by minimizing the spatial footprint of exploration activities through improved targeting efficiency, thereby reducing unnecessary ground disturbance, habitat fragmentation, vegetation clearing, and exploration-related carbon emissions.

Traditional mineral exploration approaches in Northern Nigeria have predominantly relied on geological mapping, geochemical sampling, and limited geophysical surveys. However, these conventional methods often prove insufficient for comprehensive subsurface characterization across the region's vast and often inaccessible terrains [10, 11]. The integration of gravity and magnetic data has emerged as a cost-effective and efficient approach for regional-scale mineral exploration, providing critical insights into subsurface geological structures and their potential for mineralization [3, 12, 13].

Gravity and magnetic methods have proven particularly effective in delineating basement topography, identifying structural controls on mineralization, and mapping intrusive bodies associated with mineral deposits in crystalline terrains similar to those found in Northern Nigeria [2, 14–17]. Recent studies have demonstrated the effectiveness of these methods in identifying gold mineralization in the Birnin Gwari area [18], lead-zinc deposits in the Abakaliki region, and iron ore occurrences in the Itakpe area [20]. However, the interpretation of gravity and magnetic data in complex geological environments often requires careful analytical approaches to extract meaningful geological information.

The advent of machine learning (ML) techniques has revolutionized geophysical data interpretation across various exploration domains. ML algorithms have demonstrated remarkable capabilities in pattern recognition, feature extraction, and predictive modeling when applied to geophysical datasets [2, 14, 21, 22]. In mineral exploration contexts, ML approaches have successfully been applied to identify mineralization patterns from aeromagnetic data [2], predict mineral potential from integrated geoscience datasets [2, 23], and optimize exploration targeting through multi-dimensional data analysis [14, 24].

Recent advancements in supervised learning algorithms, including Random Forest, Support Vector Machines, and Artificial Neural Networks, have shown promising results in geological mapping and mineral potential assessment [25–27]. These algorithms excel at identifying complex non-linear relationships within geophysical datasets that may not be apparent through conventional interpretation methods. Furthermore, unsupervised learning techniques such as clustering algorithms and dimensionality reduction methods have proven effective in geological domain identification and anomaly detection from gravity and magnetic data [28, 29]. The application of ensemble methods, which combine multiple ML algorithms, has gained particular attention in geoscience applications due to their ability to improve prediction accuracy and reduce overfitting [30]. Ensemble approaches have been successfully implemented in mineral prospectivity mapping, demonstrating superior performance compared to individual algorithms when dealing with complex geological datasets [31, 32].

Despite these technological advancements, the application of machine learning frameworks to

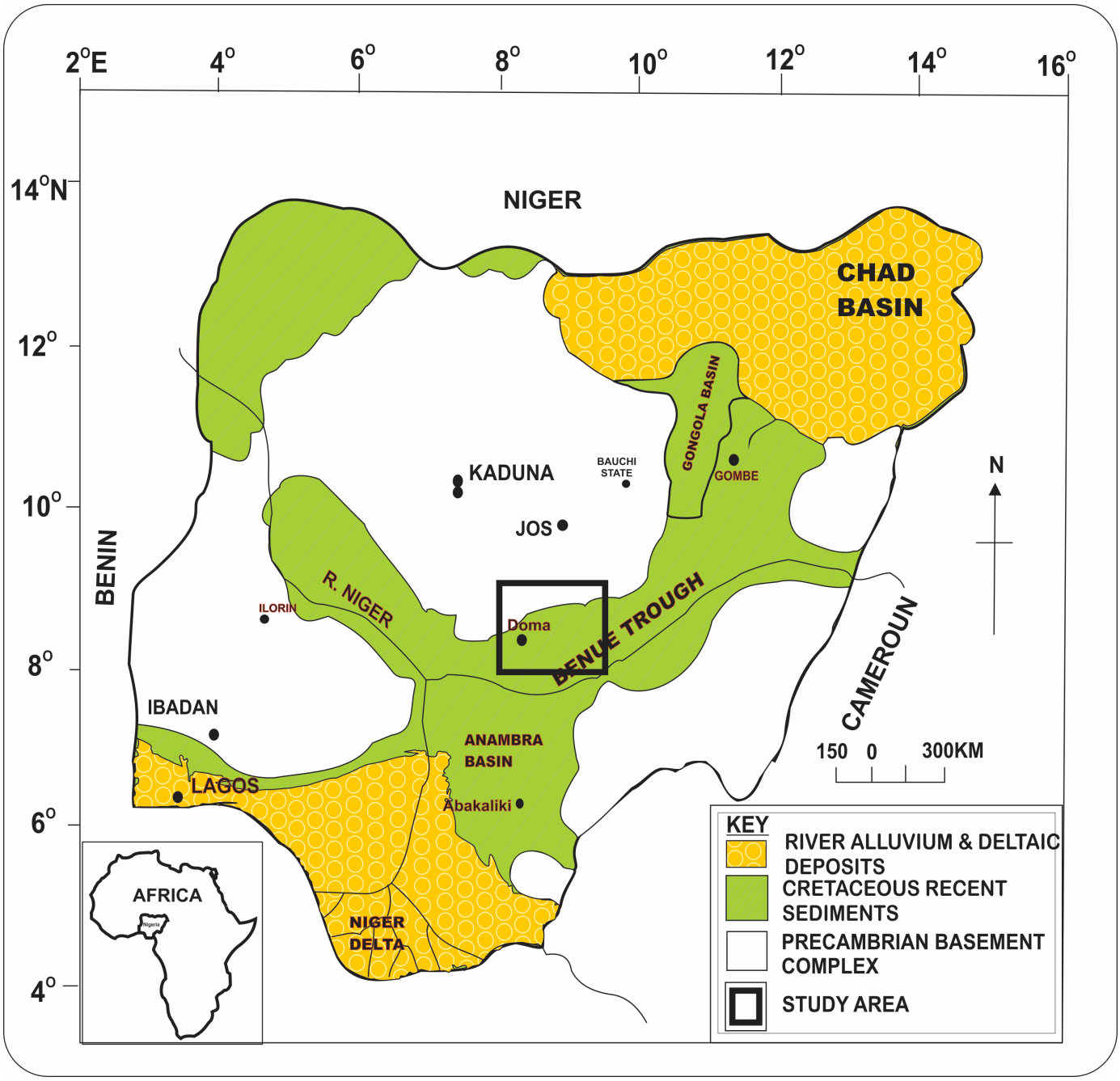

Figure 1. Map of Nigeria showing regional geology and the area covered by this study (Modified from Ugbor et al. [3]).

gravity and magnetic data interpretation for mineral exploration in Northern Nigeria remains limited. Most existing studies in the region have focused on conventional geophysical interpretation methods, with minimal integration of modern computational approaches [3, 33, 34]. This research gap presents a significant opportunity to develop innovative ML frameworks specifically tailored to the geological complexities and mineralization styles characteristic of Northern Nigeria. The successful implementation of ML frameworks in mineral exploration requires careful consideration of data preprocessing, feature engineering, algorithm selection, and validation strategies. Data preprocessing techniques, including noise reduction, grid interpolation, and derivative calculations, play crucial roles in enhancing signal quality and extracting meaningful geological information from gravity and magnetic datasets [2, 35, 36]. Feature engineering, involving the extraction of relevant geological attributes from geophysical data, significantly influences the performance of ML algorithms in geological applications [14, 37, 38].

Furthermore, the integration of multiple geophysical datasets through data fusion techniques has

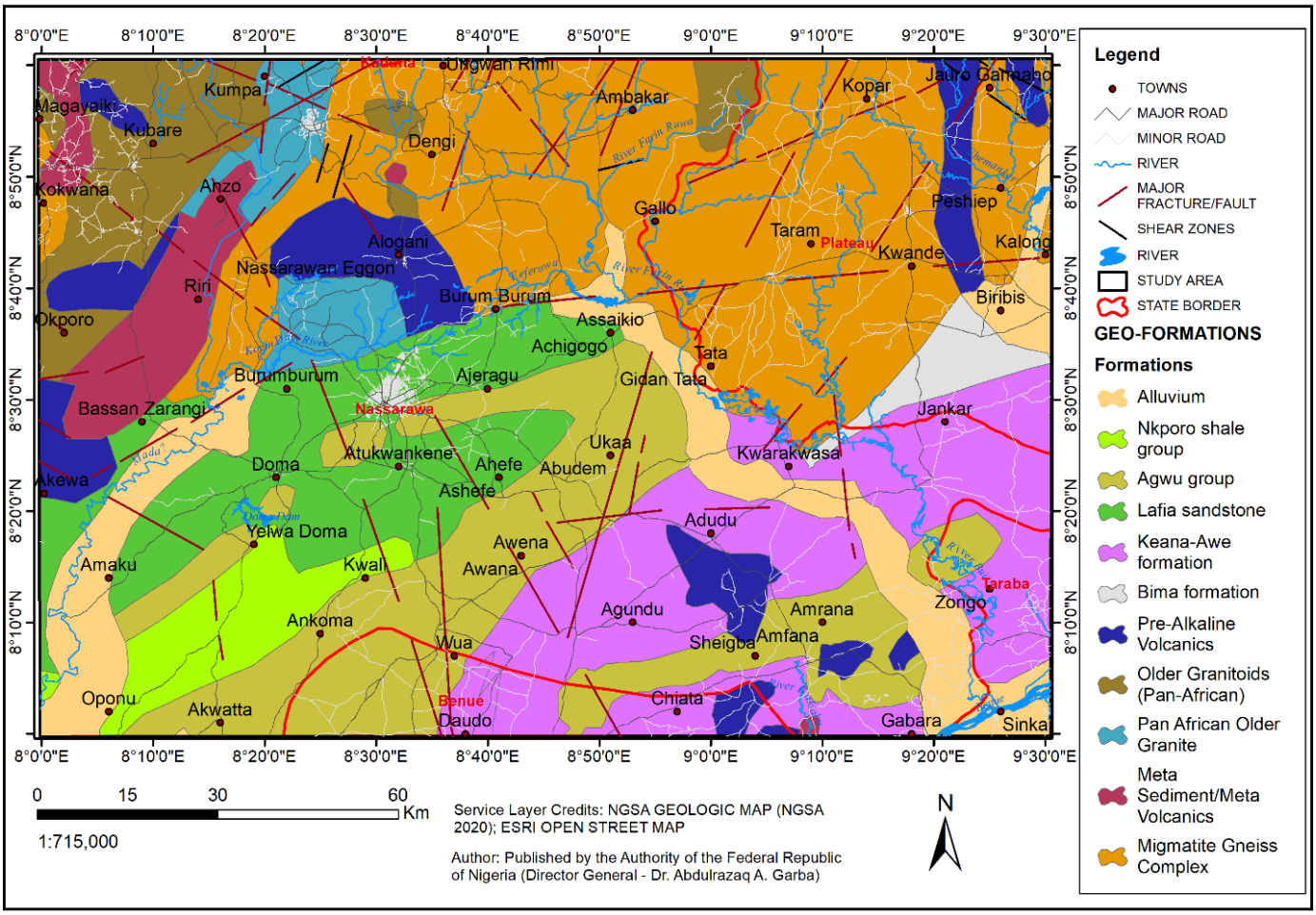

Figure 2. Geological map of the study area (extracted from Nigerian Geological Survey Agency [47]).

shown considerable promise in improving mineral exploration outcomes. Joint inversion and cooperative inversion approaches have demonstrated enhanced resolution and reduced non-uniqueness in geophysical interpretation [19, 39, 40]. The incorporation of these integration techniques within ML frameworks represents a frontier area for improving mineral exploration efficiency and accuracy.

From an environmental perspective, efficient mineral targeting through machine learning reduces the ecological impact of exploration activities by concentrating field efforts on high-probability zones, thereby minimizing vegetation clearing, soil disturbance, water resource impacts, and biodiversity disruption across broad exploration terrains [41, 42]. This precision-targeting approach aligns with sustainable mineral exploration principles and supports responsible resource development practices essential for meeting global energy transition demands while protecting environmental integrity. By reducing the area requiring ground-disturbing activities by 40-60% compared to conventional regional exploration programs, ML-guided exploration significantly decreases the environmental footprint per unit of mineral resource discovered.

This study aims to develop a comprehensive machine learning framework for investigating mineralized geologic structures in Northern Nigeria using integrated gravity and magnetic data. The research objectives include: (1) developing robust data preprocessing and feature extraction methodologies for gravity and magnetic datasets specific to the regional geological context; (2) implementing and comparing various supervised ML algorithms for geological structure identification and mineral potential mapping; and (3) establishing validation protocols for ML model performance assessment in mineral exploration applications. The methodological approaches and validation strategies developed in this study will provide valuable insights for the

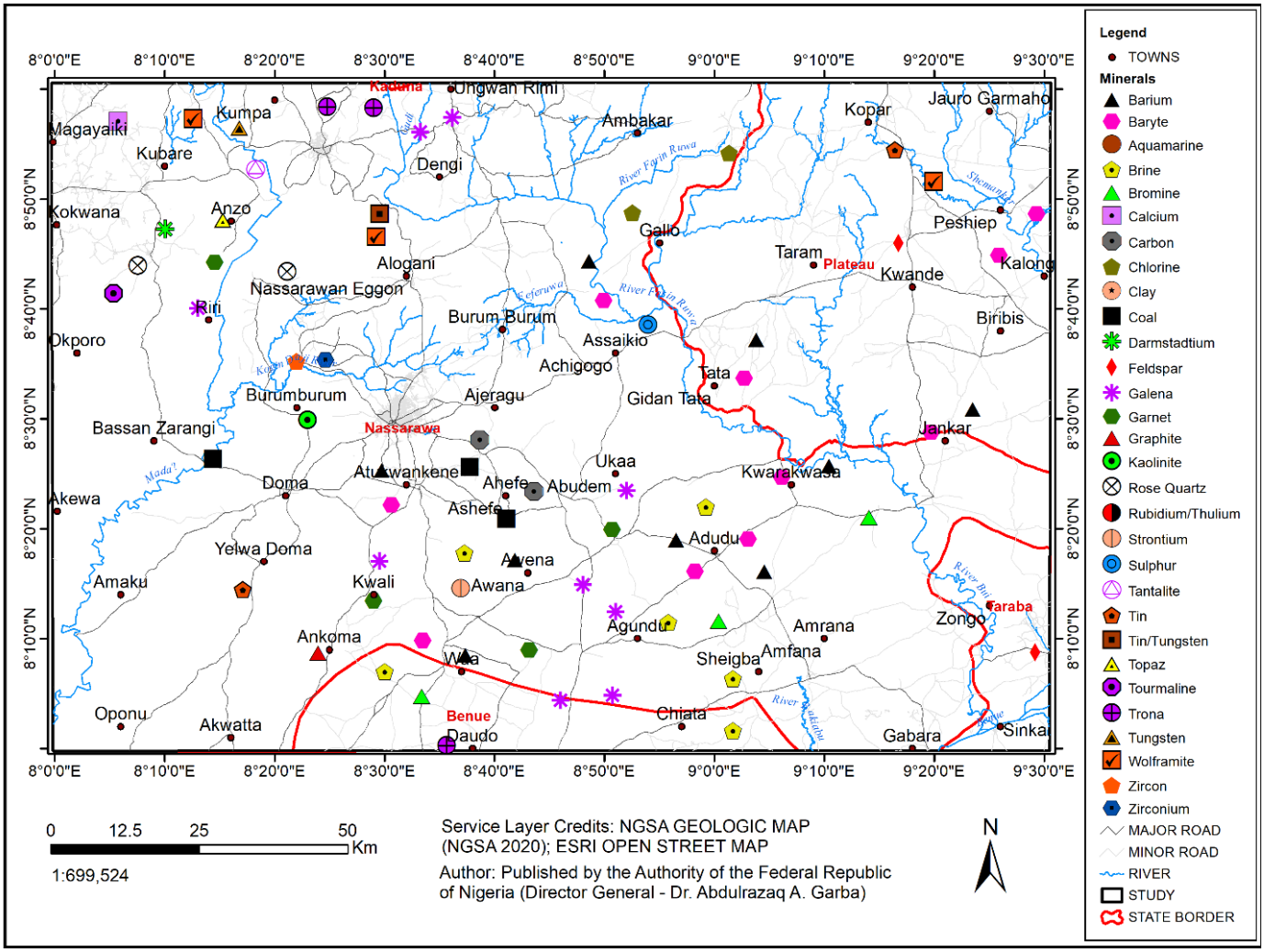

Figure 3. Mineral resources map (extracted from Mineral map of Nigerian Geological Survey Agency [51]).

geophysical exploration community and advance the integration of artificial intelligence techniques in applied geophysics.

Beyond mineral targeting, the joint gravity–magnetic machine learning workflow developed here is directly applicable to other geo-energy domains. The same multi-source geophysical integration and supervised classification approach can be adapted for geothermal reservoir delineation, where contrasts in magnetic susceptibility and density mark heat-conducting fault corridors and hydrothermal alteration zones, and for carbon dioxide (CO_2) storage site characterisation, where gravity and magnetic anomalies constrain caprock integrity, seal competence, and subsurface structural traps essential for safe sequestration [16, 43].

2 Geologic and Mineral Settings

The region of this study is situated largely within the Middle Benue Trough, a prominent intracratonic rift system that represents one of Nigeria's most significant geological features. The northern part of the study area contains the Precambrian basement complex (Figure 1) which basically divides into the Migmatite Gneiss Complex, Schist Belts, and Older Granites, forming the foundational structure of the region [44]. The Middle Benue Trough forms the central segment of the elongate, broad synclinal NE-SW trending Benue Trough, comprising areas from Makurdi through Yandev, Lafia, Obi, Jangwa to Wukari. This rift system developed through the rifting of the central West African basement complex, with initial formation commencing during the early Cretaceous period [45, 46]. The tectonic evolution of the Benue Trough reflects a complex history of continental rifting, subsidence, and intermittent compressional episodes. The main stages of tectonic evolution are well documented in the stratigraphic succession, comprising three major depositional sequences: an Albian-Cenomanian pyroclastic, paralic, shallow marine, and fluviatile sequence

corresponding to the graben and transitional tectonic stages; a Turonian-Coniacian paralic, marine and fluviatile sequence resulting from downwarping and widespread marine transgression; and a Campanian-Maastrichtian paralic, marine and fluviatile offlap sequence following the Santonian compressional deformation episode [45].

The geological succession within the study area encompasses a comprehensive record of Cretaceous sedimentation, with formations ranging from Turonian to Maastrichtian age. The study area consists of six upper Cretaceous lithogenetic formations that collectively document the complex interplay between tectonic subsidence, eustatic sea-level variations, and sediment supply dynamics. The Makurdi and Bima Formation represents one of the most significant lithostratigraphic units within the region, characterized by distinctive sedimentological and petrographic attributes (Figure 2). The formation exhibits considerable lithological heterogeneity, with alternating sequences of coarse-grained arkosic sandstones, conglomerates, and subordinate mudstone intervals. The arkosic composition indicates derivation from nearby crystalline basement sources, with limited transport distances preserving the texturally immature characteristics. Paleocurrent indicators suggest sediment transport from northeast to southwest, consistent with the regional structural grain of the underlying basement complex.

The Lafia formation is generally considered the youngest sedimentary deposits within the Mid Nigerian Benue trough, and is defined by the presence of poorly consolidated sedimentary rocks of Maastrichtian age. This is defined by an interbedded occurrence of coal seams with shale, limestone and occasional occurrence of sandstones. The formation represents the culmination of Cretaceous sedimentation within the Middle Benue Trough, deposited during a period of relative tectonic quiescence following the major Santonian deformation episode [44]. The Maastrichtian clay member of the non-fossiliferous Lafia Formation as exposed in Doma and Shabu areas provides important insights into the depositional environment and diagenetic processes affecting the youngest Cretaceous strata. The clay-rich intervals within the formation indicate deposition under low-energy, restricted marine to lacustrine conditions, with periodic emergence evidenced by coal seam development. The Lafia Formation exhibits significant economic potential, particularly regarding coal resources and industrial clay deposits. Selected coals from Chikila (CHK), Lafia Obi (LFB) and Okaba (OKB) in the Benue Trough have been extensively studied for their geomineralogical and thermal characteristics (Figure 3), revealing substantial reserves of sub-bituminous to bituminous coal suitable for energy production and industrial applications [48–50].

The structural architecture of the Lafia-Makurdi region reflects the complex tectonic history of the Benue Trough system. Lineament analysis obtained from geophysical data shows linear structures that trend in the NW-SE and E-W directions (Figure 2) which could be interpreted as veins that are host for minerals within the area, predominant around Kwalla [10]. These structural trends represent reactivated basement fabrics and syn-sedimentary fault systems that controlled both sedimentation patterns and post-depositional fluid migration [52]. The dominant structural grain exhibits a NE-SW orientation, parallel to the regional Benue Trough axis, with subsidiary NW-SE and E-W trending structures representing conjugate fracture systems developed during episodes of transpressional deformation [53]. These structural elements are particularly significant for mineral exploration, as they commonly serve as conduits for hydrothermal fluid circulation and sites of ore deposition [1–3, 19, 54].

The geological setting of the Lafia-Makurdi region provides favorable conditions for various types of mineralization (Figure 3). The combination of basement-hosted structures, hydrothermal alteration associated with igneous intrusions, and favorable host rocks within both basement and sedimentary sequences creates diverse metallogenic environments. Coal resources within the Lafia Formation represent the most significant economic deposits currently identified, with substantial reserves of varying rank and quality. Industrial minerals, including clays suitable for ceramic and construction applications, are abundant within the Maastrichtian sequences. The structural framework, particularly the intersection of NW-SE and E-W trending lineaments with the main NE-SW regional trend, provides potential loci for hydrothermal mineralization. These structural intersections are commonly associated with base metal, precious metal, and rare earth element concentrations in similar geological environments globally [50, 55, 56].

3 Methodology

3.1 Magnetic Data

High-resolution airborne magnetic data for this investigation was obtained from the Nigerian Geological Survey Agency (NGSA) archives, representing comprehensive coverage from the national aerophysical mapping program executed by Fugro Airborne Surveys during 2005-2009. The magnetic survey specifications included systematic data collection along parallel flight traverses spaced at 500 m intervals, with aircraft maintaining a consistent terrain clearance of approximately 80 m above ground level. Flight path orientation was established in a northwest-southeast direction to optimize magnetic anomaly resolution perpendicular to the dominant regional geological strike.

Data acquisition procedures incorporated standard magnetic surveying protocols with continuous recording of total magnetic intensity measurements using cesium vapor magnetometers. The resulting geophysical database was delivered in gridded format with 30-arc-minute spatial resolution tiles, facilitating regional-scale interpretation and subsequent detailed analysis of localized magnetic signatures. Comprehensive data processing workflows were implemented to enhance signal quality and geological interpretation capability. Diurnal magnetic field variations were systematically corrected using reference station measurements recorded throughout each survey flight, while regional magnetic field components were removed through application of the International Geomagnetic Reference Field (IGRF) model appropriate for the survey period.

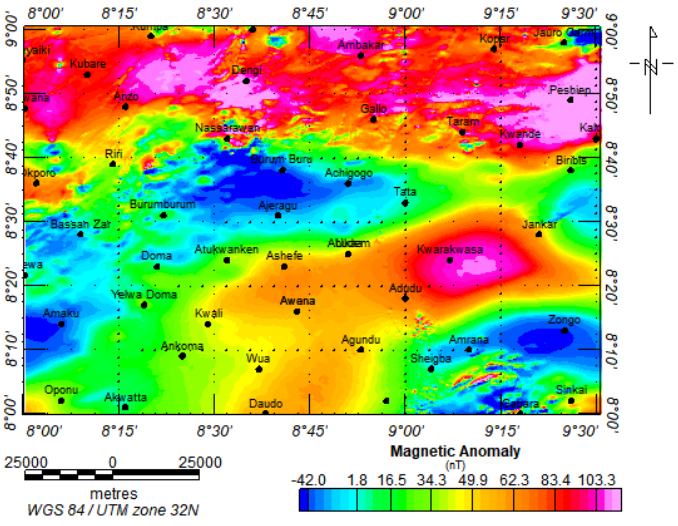

Figure 4. Magnetic anomaly map of the study area.

Magnetic pole reduction procedures were applied to transform the observed magnetic anomalies to equivalent anomalies that would be observed at the magnetic equator, thereby simplifying interpretation by centering magnetic anomalies directly over their causative geological sources. This transformation utilized established methodologies following the theoretical framework outlined by Ugbor et al. [3], Abraham et al. [43], Roest et al. [57], and Nabighian [58]. The reduction calculations incorporated local geomagnetic parameters specific to the study region, including magnetic declination of -2.15° and inclination of -13.91°, implemented through frequency-domain fast Fourier transform algorithms. The final processed aeromagnetic grid exhibits enhanced geological resolution with suppressed directional bias effects, providing optimal conditions for subsequent machine learning analysis and geological feature extraction. The resulting magnetic intensity map demonstrates clear delineation of subsurface geological structures and is presented in Figure 4 for detailed examination.

3.2 Gravity Data

The gravitational field dataset employed in this investigation was derived from advanced satellite altimetry measurements, incorporating state-of-the-art radar altimetry missions including CryoSat-2 and Jason-1 platforms [59, 60]. These contemporary satellite systems feature enhanced track density coverage and advanced radar instrumentation, delivering substantial improvements in range measurement precision that directly enhance the resolution and accuracy of derived gravity field models. The enhanced technological capabilities of modern altimetry systems provide markedly superior data quality relative to conventional marine gravity observations collected through traditional shipborne surveys [59]. The integration of multiple satellite altimetry missions with comprehensive ground-based gravity observations has facilitated the development of high-fidelity gravitational models suitable for detailed subsurface geological investigations [61].

Contemporary research applications have demonstrated the exceptional capability of satellite-derived gravitational data for mapping complex tectonic and structural geological features across diverse geological terrains [62, 63]. For the present study, Free-Air gravitational anomaly data and corresponding digital elevation models were acquired to provide comprehensive coverage of the study

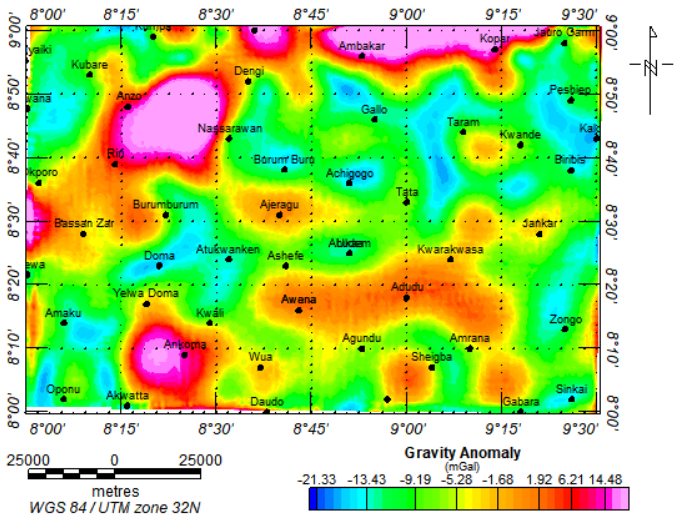

Figure 5. Magnetic anomaly map of the study area.

region. Initial data processing involved interpolation to a uniform 1km grid spacing using minimum curvature algorithms, ensuring optimal spatial resolution while maintaining data integrity across the survey area. Complete Bouguer anomaly calculations were subsequently performed using standard crustal density assumptions of 2.67 g/cm^3, representing typical continental crustal compositions appropriate for the geological setting of the investigation area. Regional-residual field separation procedures were implemented using least-squares polynomial fitting techniques to distinguish long-wavelength regional gravity trends from shorter-wavelength anomalies associated with localized geological features. The final processed residual gravity dataset was interpolated to a consistent 1km grid resolution, producing the comprehensive gravity anomaly map illustrated in Figure 5.

3.3 Analytic signal (AS) Evaluation

The analytic signal transformation represents a mathematical approach that integrates both magnitude and phase characteristics derived from magnetic field measurements, thereby enhancing the detection and delineation of subtle magnetic variations associated with lithological contrasts and structural geological boundaries [1, 64, 65]. This processing technique demonstrates particular effectiveness in geologically complex environments where conventional magnetic interpretation approaches encounter limitations due to signal interference phenomena and variable magnetization vector orientations [3, 58]. The analytic signal function (Equation 1), while not representing a directly observable physical parameter, serves as a powerful interpretive tool in geophysical analysis due to its inherent independence from geomagnetic field orientation and magnetization direction influences [66, 67]. This fundamental property ensures that geological bodies exhibiting comparable geometric configurations generate equivalent analytic signal responses, regardless of their geographic location or magnetic latitude positioning.

The Analytic Signal is given by Equation (1):

$$A(x,y) = \sqrt{\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2} \qquad (1)$$

where T is the observed field at x and y.

Furthermore, the analytic signal technique produces symmetric response patterns with peak amplitudes positioned directly above the edges of extensive geological bodies or centrally located over more restricted structures, thereby providing precise spatial information regarding subsurface geological feature positioning [35]. This geometric relationship between analytic signal maxima and geological source boundaries facilitates accurate structural mapping and enhances the reliability of geological interpretation derived from magnetic datasets. The implementation of analytic signal processing in this investigation enables enhanced resolution of geological contacts, improved delineation of intrusive body margins, and more accurate identification of structural features that may control mineralization processes. The resulting analytic signal maps provide optimal input datasets for machine learning algorithms, as they present geological information in a standardized format that facilitates automated pattern recognition and feature extraction procedures [2].

3.4 Source Parameter Imaging (SPI)

The Source Parameter Imaging methodology represents a depth estimation approach that directly processes gridded magnetic intensity data without necessitating magnetic pole reduction transformations. This fundamental advantage eliminates the computational complexities associated with magnetic inclination and declination corrections, thereby simplifying the analytical workflow while maintaining robust depth determination capabilities. The method's independence from geomagnetic coordinate transformations provides particular

utility in low-latitude magnetic surveys where pole reduction procedures may introduce interpretation uncertainties.

SPI assumes a step-type source model [12, 68]. For a step, the following formula (Equation 2) holds [3]:

$$\text{Depth} = \frac{1}{K_{\max}} \tag{2}$$

where $K_{\max}$ is the peak value of the local wavenumber K over the step source.

Quantitative validation studies incorporating borehole control data have established that Source Parameter Imaging consistently achieves depth estimation accuracies within $\pm$20% of known subsurface targets. This performance standard compares favorably with established depth estimation techniques such as Euler deconvolution methods [2, 3], while offering distinct computational advantages through enhanced spatial resolution and continuity of solution distribution. The superior density of SPI-derived depth estimates facilitates more comprehensive subsurface structural interpretation compared to conventional automated depth estimation approaches. Implementation of SPI analysis across the study area generated high-resolution depth-to-source distributions that reveal detailed subsurface structural architecture previously unresolved through conventional interpretation methods. The comprehensive spatial coverage achieved through SPI processing provides continuous depth mapping capabilities that significantly enhance geological understanding of subsurface complexity patterns.

3.5 Machine Learning (ML) Algorithms

3.5.1 Unsupervised Clustering Analysis

K-means clustering is applied to identify natural groupings within the geophysical data using longitude, latitude, and analytic signal magnitude as input features. The clustering algorithm employs a default of three clusters (k=3) with random initialization seed set to 42 for reproducibility. This unsupervised approach aids in understanding the inherent spatial patterns within the geophysical data and provides additional contextual information for supervised classification.

3.5.2 Supervised Machine Learning Framework

Feature Engineering. The final feature vector for supervised learning incorporates six variables: geographic coordinates (longitude, latitude), magnetic anomaly amplitude, gravity measurements, analytic signal magnitude, and depth information from SPI. This multi-parameter approach leverages the complementary nature of magnetic [2] and gravimetric methods for geological characterization.

Model Selection and Training. Five distinct machine learning algorithms are implemented [2, 14] and compared for geological structure classification:

- Logistic Regression: A linear probabilistic classifier with maximum iterations set to 1,000 to ensure convergence,
- Decision Tree Classifier: A non-parametric method using recursive binary splitting,
- Support Vector Machine (SVM): Implemented with radial basis function kernel and probability estimates enabled,
- Random Forest: An ensemble method combining multiple decision trees with bootstrap aggregating,
- Gradient Boosting Classifier: A sequential ensemble technique using adaptive boosting principles.

All models utilize identical random seeds (42) for reproducible results and consistent cross-validation procedures.

3.5.3 Model Evaluation and Validation

The dataset is partitioned using stratified sampling with an 80:20 training-to-testing ratio [2]. Model performance is assessed using multiple metrics:

- Classification Accuracy: Overall correct prediction rate,
- Precision: True positive rate for each class (macro-averaged),
- Recall: Sensitivity or true positive rate (macro-averaged),
- F1-Score: Harmonic mean of precision and recall (macro-averaged),
- Area Under the ROC Curve (AUC): Discrimination capability assessment,
- Confusion Matrix: Detailed classification performance breakdown.

Receiver Operating Characteristic (ROC) curves are generated for models supporting probabilistic predictions, enabling comprehensive performance comparison and optimal threshold selection.

3.5.4 *Model Optimization and Selection*

The best-performing model is identified using AUC score as the primary selection criterion, as this metric provides robust performance assessment for potentially imbalanced geological datasets [14]. Hyperparameter optimization could be implemented through grid search or random search methodologies, though default parameters are utilized in the baseline implementation.

3.5.5 *Prediction and Validation*

The optimal model generates predictions for the entire dataset, enabling spatial mapping of geological structures. Prediction accuracy is evaluated by comparing model outputs with ground-truth classifications. Misclassified samples are isolated and analyzed separately to identify potential systematic errors or challenging geological scenarios requiring further investigation.

All computational analyses were performed using a Python script written on Python 3.8 with scientific computing libraries including NumPy, Pandas, Scikit-learn, SciPy, and Matplotlib. The methodology provides a robust, reproducible framework for automated geological structure classification using integrated geophysical datasets and machine learning techniques.

3.6 Feature Ablation Experiments

To evaluate the independent predictive contribution of each geophysical parameter, a comprehensive ablation study was introduced. Three experiments were performed:

1. **Leave-One-Out Ablation:** Models were retrained iteratively while removing one feature at a time (AS, magnetic anomaly, gravity, or SPI depth). Changes in ROC-AUC, precision, and F1 were recorded to quantify the dependence of the model on each feature.
2. **Feature-Set Comparison Ablation:** Three separate models were trained using:
   (a) magnetic anomaly + analytic signal only,
   (b) gravity + SPI depth only,
   (c) all features combined.

   This allowed evaluation of whether magnetic-only predictions outperform gravity-only predictions and whether the combination provides synergistic benefits.
3. **Cross-validated ROC Stability Analysis:** For each feature set, 10-fold cross-validated ROC-AUC values were computed and averaged to assess robustness and minimize sampling bias.

3.7 Correlation and Feature-Importance Analysis

A full feature-correlation matrix (Pearson) was computed to determine whether any predictor was collinear with the target or with each other. Correlation values are presented as a heatmap. This provided an explicit statistical basis for evaluating inherent relationships between analytic signal, gravity anomalies, depth, and mineralization. Model-based Random Forest feature importance and permutation importance were also computed. These metrics quantify the dependency of model predictions on each geophysical feature. This dual evaluation revealed which physical properties exert the strongest control on mineralization signatures in the dataset. A summary of the processing steps taken in this study is presented in the flowchart below (Figure 6):

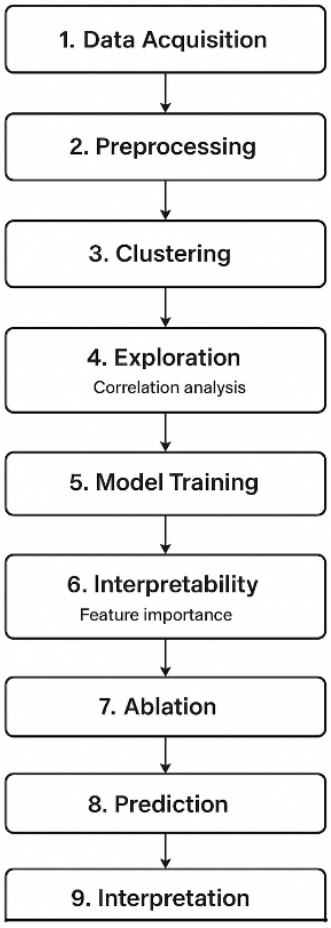

Figure 6. A flowchart presentation of the processing steps followed in this study.

4 Results

The analytic signal transformation results are presented in Figure 7, demonstrating the enhanced interpretive capabilities of this magnetic data processing technique. The analytic signal response across the investigation area constituted a fundamental component of the comprehensive feature extraction framework implemented within the machine learning analytical pipeline. This data transformation technique converts raw magnetic field measurements into interpretively enhanced datasets that accentuate geological boundaries [69, 70] and structural discontinuities present within the subsurface geological architecture of the study region. The feature engineering implementation of analytic signal analysis effectively transforms magnetic field data into a representation optimized for identifying geological transitions, including fault systems, lithological contacts, and mineralized structural zones. This mathematical transformation serves as a geological proxy, facilitating the systematic delineation of subsurface structural elements through automated pattern recognition algorithms integrated within the machine learning framework.

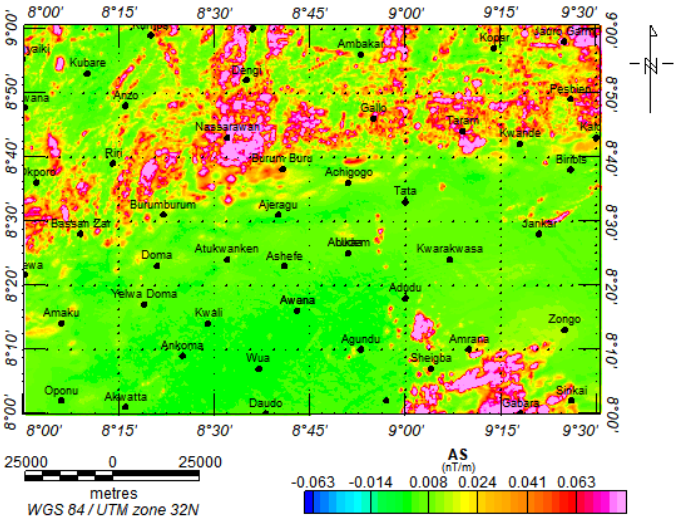

Figure 7. Analytic signal map of the study area.

The computational results from SPI analysis is presented in Figure 8, demonstrating the detailed structural information extracted through this processing approach. The derived depth parameters serve as critical input variables for machine learning model development, providing quantitative subsurface structural characterization that enhances automated geological feature recognition capabilities. Integration of SPI-derived depth information as supplementary feature vectors will significantly improve the discriminatory power of machine learning algorithms applied to the geological structure classification tasks. These depth-related attributes provided essential three-dimensional geological context that augments the interpretive capability of surface magnetic anomaly patterns, thereby enhancing the overall effectiveness of automated mineral exploration targeting procedures.

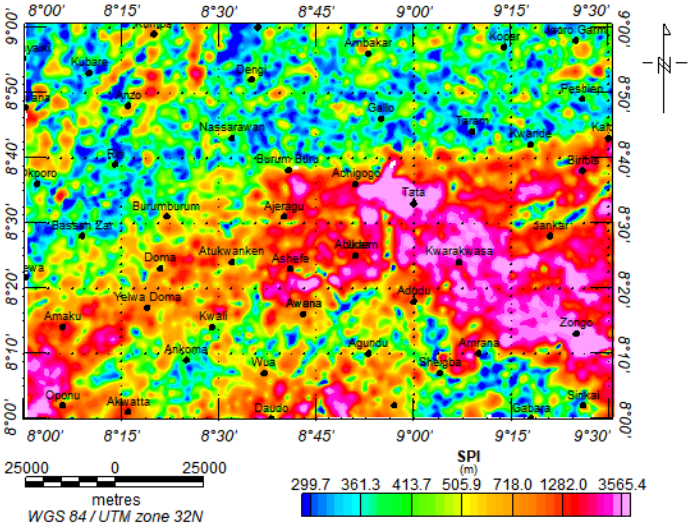

Figure 8. SPI results for the study region.

Comprehensive performance metrics for all evaluated machine learning algorithms including a cross-validated ROC stability analysis are documented in Table 1, providing detailed comparative analysis of classification accuracy, precision, recall, and F1-score statistics across the tested algorithmic approaches. The ROC curves, along with corresponding AUC calculations for each implemented model, are graphically illustrated in Figure 9. These diagnostic plots demonstrate the trade-off relationships between true positive and false positive rates across varying classification thresholds, enabling quantitative assessment of model discriminatory capabilities for mineralized structure identification.

Using the full set of geophysical predictors (magnetic anomaly, analytic signal, gravity anomaly, and SPI-derived depth), all five classifiers achieved high predictive performance (Figure 9), with ROC-AUC values exceeding 0.94 for most models. The Random Forest classifier produced the best overall performance, achieving the highest ROC-AUC (0.9776), accuracy (0.9540), and macro-F1 (0.8497) (Table 1). Gradient Boosting followed closely with a ROC-AUC of 0.9674. Support Vector Machine (SVM) performed competitively (ROC-AUC = 0.9628), while Logistic

Table 1. Model accuracy comparison.

Parameters	Random Forest Model	Gradient Boosting Model	Logistic Regressions	SVM	Decision Tree
Accuracy	0.954	0.946	0.935	0.940	0.933
Precision	0.889	0.946	0.857	0.876	0.803
Recall	0.819	0.788	0.707	0.728	0.792
F1 Score	0.849	0.821	0.757	0.779	0.797
ROC-AUC Score	0.978	0.967	0.948	0.963	0.792
CV ROC-AUC	0.977	0.967	0.948	0.962	0.796

*Best model: Random Forest

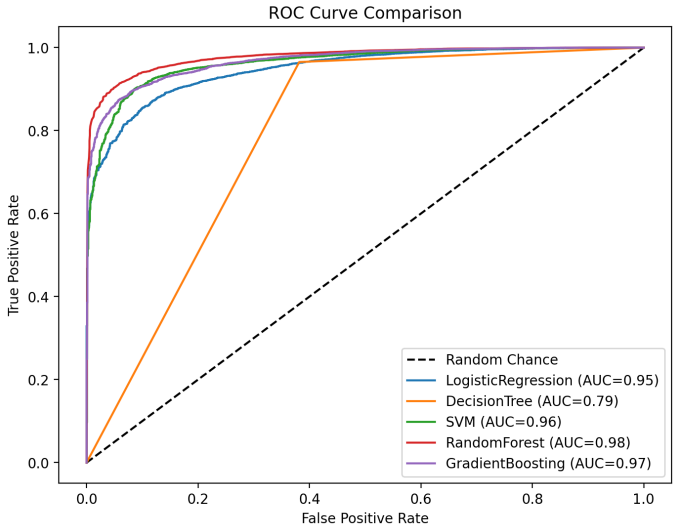

Figure 9. ROC curve and AUC for the evaluated models.

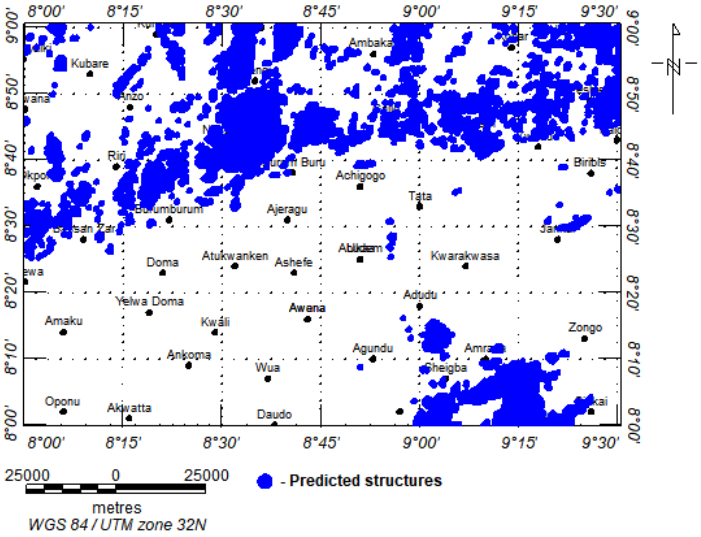

Figure 10. Predicted mineralized structures from the RF model evaluations.

Regression and Decision Tree models demonstrated acceptable but lower predictive capacity (ROC-AUC = 0.9480 and 0.7918 respectively).

Ten-fold cross-validation confirmed the stability of the ensemble models, with Random Forest yielding the smallest standard deviation (0.00092) in CV ROC-AUC. These results indicate that mineralization patterns in the study area are nonlinear and better captured by ensemble approaches that exploit interactions among features.

The per-point predictions (Figure 10) reveal that the combined model produces spatially coherent mineralization probabilities. Predicted mineralized pixels tend to cluster along major structural corridors aligned NW–SE and NE–SW, consistent with known fracture and intrusive trends in the region.

To assess the independent contribution of each geophysical feature group, three models were trained using different subsets of inputs. Results from the feature-set ablation (Figure 11) show clear differences in predictive performance. The Magnetic

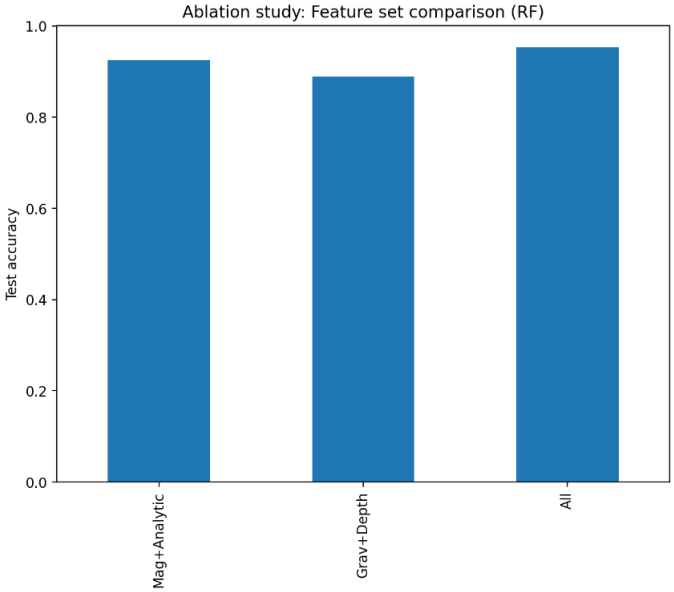

Figure 11. Plotted feature – set ablation evaluation of predictive contributions of each geophysical parameter to the model.

Anomaly + Analytic Signal Only configuration produced high ROC-AUC values, confirming that magnetic-derived attributes are strong predictors of structurally controlled mineralization in the region. The Gravity + SPI-Depth Only models trained solely on gravity and SPI depth delivered moderate but meaningful prediction performance. Although their ROC-AUC was lower than magnetic-based models, they successfully identified several zones where magnetic signatures are weak or suppressed, indicating complementary sensitivity to deep-seated density contrasts and structural roots. The All Features Combined model consistently outperformed both individual subsets. This demonstrates that gravity and depth information improve the predictive landscape, particularly in magnetically subdued areas, thereby reducing false positives associated with AS hotspots and increasing true detections in deeper or altered zones. Overall, the ablation study confirms that although analytic signal is the strongest individual predictor, the best discrimination is achieved only when all geophysical attributes are integrated.

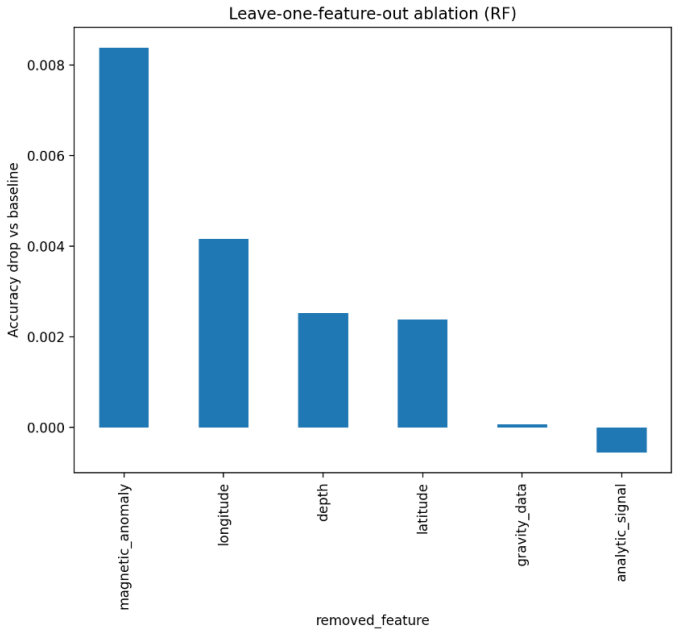

Figure 12. Plot of the leave-one-out ablation analysis performed on the geophysical parameters used in the modeling.

The leave-one-out ablation analysis (Figure 12) provides a more granular understanding of the relative importance of each feature. Removing AS resulted in the largest reduction in ROC-AUC and F1 score. This confirms its dominant contribution, as expected in magnetically responsive basement terrains. Removing magnetic anomaly produced a moderate drop in performance, indicating that the magnetic dataset contributes independent shallow-structure information not fully captured by AS alone. Removing gravity anomaly caused noticeable degradation in model stability and increased misclassification in low-magnetic areas. This supports the idea that the gravitational field helps constrain deeper lithological boundaries and density-controlled mineralization. Removing SPI depth also reduced predictive performance, confirming its value in delineating the depth-to-source variations that influence mineral deposition.

Collectively, the leave-one-out results show that all four features contribute meaningfully, though to different extents, and that the model's success cannot be attributed solely to high AS amplitudes.

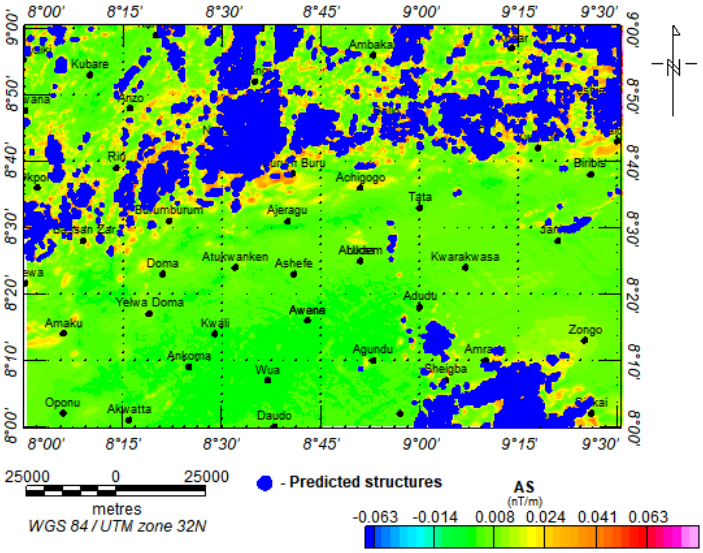

Figure 13. A comparison of results from the conventional AS evaluations (Figure 7) and predicted subsurface mineralized structures from the best model (RF).

Figure 13 displays a comparison plot of the predicted structures with the computed AS results.

A comprehensive comparison is made between conventional analytic signal response and the RF model predictions of subsurface geological structures. The Pearson correlation heatmap is presented in Figure 14. Result from K-means clustering analysis is shown in Figure 15.

5 Discussion

5.1 General Discussion

Across all analyses, AS (Figure 7) remains the single strongest predictor, which is consistent with the physics of the method. AS enhances edges, faults,

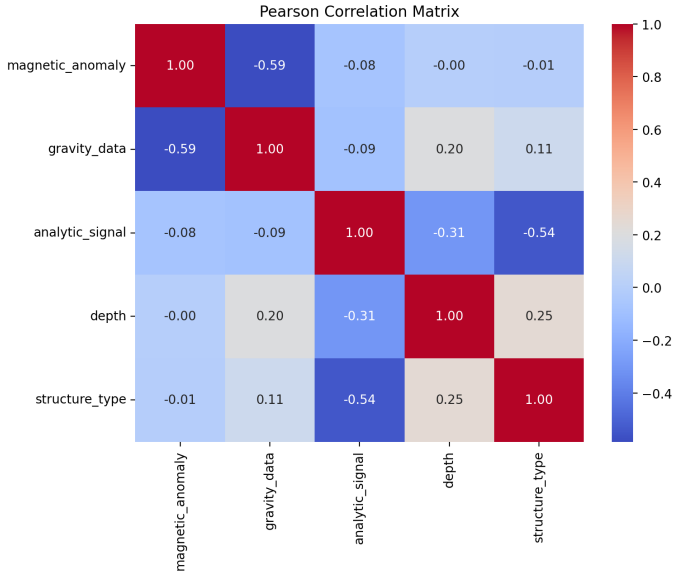

Figure 14. The Pearson correlation heatmap presenting the feature correlation analysis.

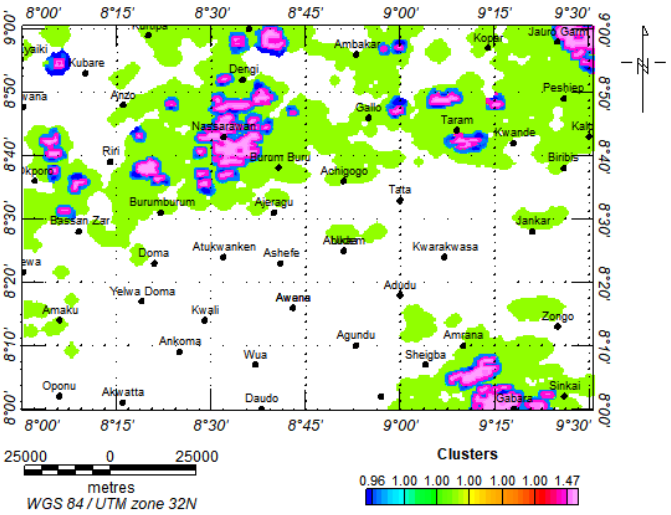

Figure 15. Unsupervised clustering analysis (K-means clustering, k=3).

and intrusive contacts (features known to host the mineral occurrences in the region) [67]. However, reliance on AS alone would bias predictions toward near-surface and magnetically strong zones. The inclusion of gravity and SPI depth (Figure 8) therefore strengthens geological plausibility by capturing; deep structural roots that do not express strong magnetic signatures, density contrasts associated with alteration zones, thickening of regolith or basin-fill that suppresses magnetic response, and intrusive formations detectable in both gravity and depth inversion. This integrated behaviour aligns with regional geological expectations and demonstrates that the final model is not merely reproducing AS highs but is responding to a broader suite of geophysical expressions of mineralization.

The multi-model evaluation (Logistic Regression, Decision Tree, SVM, Random Forest, Gradient Boosting) shows a consistent pattern (Table 1). Eensemble learners (Random Forest and Gradient Boosting) performed best overall, achieving the highest ROC-AUC (0.98 and 0.97 respectively) and the most stable cross-validated estimates. These models capture nonlinear interactions among features, supporting our deduction that mineralization in the study area is controlled by a combination of magnetic contrasts, structural depth variations, and subtle gravity signatures rather than a single dominant parameter. The superior performance of Random Forest in particular reinforces the importance of feature interactions. The ROC curve analysis presented in Figure 9 demonstrates the comparative performance of five machine learning algorithms in distinguishing between mineralized and non-mineralized geological structures within the study area. All evaluated models exhibit superior discriminatory capability compared to random classification, as evidenced by their substantial deviation above the diagonal reference line representing random chance performance.

The Random Forest algorithm achieved the highest classification performance with an AUC score of 0.98, indicating exceptional ability to correctly rank mineralized locations above non-mineralized sites across all probability thresholds. This near-perfect AUC value demonstrates that the Random Forest model successfully captures the complex nonlinear relationships between geophysical features and mineralization occurrence patterns, making it reliable for exploration targeting applications.

Gradient Boosting Machine demonstrated competitive performance with an AUC of 0.97, positioning it as the second-best performing algorithm. The marginally lower performance compared to Random Forest suggests that while both ensemble methods effectively model the underlying geological relationships, Random Forest's bagging approach with feature randomization provides slight advantages over the sequential boosting strategy for this particular geological dataset. Logistic Regression achieved a robust AUC score of 0.95, demonstrating that even linear classification approaches can effectively differentiate mineralized from barren areas when appropriate feature engineering is applied. This performance indicates the presence of strong linear relationships between certain geophysical attributes

and mineralization occurrence, validating the geological relevance of the selected input features. Support Vector Machine also delivered good performance with an AUC of 0.96, confirming the method's capability to identify complex decision boundaries in high-dimensional feature space. The slightly lower performance relative to ensemble methods suggests that the geological relationships may benefit from the averaging effects and overfitting resistance inherent in Random Forest and Gradient Boosting approaches [2, 14]. Decision Tree exhibited the most modest performance among the evaluated algorithms with an AUC of 0.79, though still maintaining substantial improvement over random classification. The lower performance likely reflects the algorithm's tendency toward overfitting and its sensitivity to small variations in training data, particularly challenging given the geological complexity and potential noise in geophysical datasets we used.

The ROC curves reveal distinct performance characteristics among the algorithmic approaches. The ensemble methods (Random Forest and Gradient Boosting) demonstrate superior performance across all threshold settings, with particularly strong performance in the high-sensitivity region of the curve. This characteristic is especially valuable for mineral exploration applications where maximizing the detection of true mineralized zones while minimizing false positives is critically important for exploration cost-effectiveness [2]. The steep initial rise of the Random Forest curve indicate excellent performance at high-probability thresholds, suggesting the model can reliably identify the most prospective exploration targets. Conversely, the more gradual slope of the Decision Tree curve indicates reduced discriminatory power, particularly for moderate-probability classifications where exploration decision-making becomes more challenging. The exceptional performance of the Random Forest algorithm (AUC = 0.98) provides some confidence in its ability to identify previously unrecognized mineralized structures within the study area. The model's high AUC score indicates that approximately 98% of randomly selected mineralized locations will receive higher probability scores than randomly selected non-mineralized locations, representing exceptional predictive reliability for exploration targeting.

The consistently strong performance across multiple algorithmic approaches validates the quality and geological relevance of the integrated geophysical feature set derived from gravity and magnetic data. The fact that all models significantly outperform random classification confirms that the geophysical signatures contain genuine geological information capable of distinguishing mineralization-favorable from unfavorable structural settings. Based on the ROC analysis, the Random Forest algorithm emerges as the optimal choice for operational mineral exploration targeting within the study region. Its superior AUC performance, combined with the inherent resistance to overfitting and ability to handle complex feature interactions, makes it the most suitable approach for generating reliable exploration probability map. Therefore we adopted the Random Forest model for the prediction evaluations of the subsurface mineralized geologic structures within the study area. The spatial distribution of predicted geological structures, shown as blue polygons (Figure 10), demonstrates remarkable correlation with zones of elevated analytic signal amplitude, particularly in the northern and eastern regions of Ambakar, Gallo and Nassarawan Eggon. These regions are largely located on the Migmatite gneiss complex of the study area. This correspondence validates the ability of machine learning model to recognize and extrapolate the geophysical signatures associated with geological structural boundaries that are conventionally identified through analytic signal analysis.

The ablation experiments provide clear insight into the relative contribution of the different geophysical datasets (magnetic anomaly, analytic signal, gravity, and SPI-derived depth) to the prediction of mineralized structures (Figure 11). Three feature groups were evaluated: (1) Magnetic Anomaly + Analytic Signal, (2) Gravity + Depth, and (3) All Features Combined. Their respective test AUC values were 0.881, 0.861, and 0.977, demonstrating that although analytic signal and magnetic anomaly contain the strongest individual discriminatory power, the integration of all geophysical inputs produces a substantially more accurate and robust model.

The first ablation set, comprising magnetic anomaly and analytic signal only, achieved an AUC of 0.881, confirming that high analytic-signal responses correlated with mapped mineralized structures. However, this experiment also reveals that analytic signal alone does not fully account for the predictive skill of the model; performance remains significantly below that obtained from the full feature set

(Figure 11). The second ablation set, using only gravity and SPI-derived depth, achieved an AUC of 0.861, indicating that these features contain independent structural and lithological information relevant to mineralization. While their predictive capacity is lower than magnetic/analytic-signal features, they clearly provided non-redundant information, as the model retains meaningful discrimination capability even without magnetic inputs. This shows that gravity and depth features encode geologically relevant signals, likely related to density variations associated with faults, alteration zones, or lithologic contacts in the study area.

The most compelling evidence emerges from the full feature model, which achieved the highest performance across all metrics (AUC = 0.977, accuracy = 0.953). This improvement of +0.096 AUC over the magnetic+analytic subset, and +0.116 AUC over the gravity+depth subset, demonstrates strong complementarity among the geophysical parameters. In other words, the model benefits not from a single dominant signal but from the integration of multiple physical property contrasts. This synergy supports our interpretation that mineralization in the study area is controlled by a combination of structural complexity, magnetic susceptibility contrasts, and density-related features, each captured differently by the gravity, magnetic anomaly, analytic signal, and SPI depth components. The leave-one-feature-out analysis (Figure 12) further reinforces this conclusion. Removing any single feature resulted in a measurable drop in accuracy, with the largest decline occurring when magnetic anomaly was removed (accuracy drop = 0.0084). Notably, even removing gravity or depth still produced a small but consistent reduction in performance, confirming their additive value despite being less dominant. Collectively, these results demonstrates that model predictions do not arise from analytic signal alone; rather, they reflect a multi-parameter geophysical signature of mineralized structures.

The northern sector of the study area (Figure 13), encompassing the regions around Kubare, Amabie, and extending eastward toward Birbis, exhibits extensive zones of high analytic signal amplitude (orange to red colors) that correspond closely with dense clusters of machine learning-predicted structures. This spatial concordance indicates that the Random Forest algorithm has successfully learned to identify the complex geophysical signatures associated with major structural discontinuities and lithological boundaries characteristic of basement complex terrains.

In the central portion of the map, around Lafia, Doma, and Atukwanken, the analytic signal reveals more moderate amplitude variations (green to yellow colors) (Figure 13), with the machine learning predictions identifying discrete structural zones that align with subtle linear features in the analytic signal pattern. This demonstrates the capability of model to detect structurally-controlled features that may be less apparent through conventional interpretation methods, particularly in areas where magnetic signature contrasts are more subdued. The strong spatial correlation between conventional analytic signal analysis and machine learning predictions provides robust validation of the Random Forest model's geological interpretation capabilities [2, 14]. However, the machine learning approach offers significant advantages through its ability to integrate multiple geophysical parameters simultaneously, rather than relying solely on magnetic gradient analysis.

Notably, the predicted structures extend into regions where analytic signal amplitudes are relatively low (blue to green areas) (Figure 13), particularly in the southwestern and south-central portions of the study area around Ankoma, Wua, and Daudo. This extension capability demonstrates the capacity of the model to identify potentially mineralized structures in areas where conventional magnetic interpretation might overlook subtle but geologically significant features. The machine learning predictions reveal structural complexity patterns that complement and extend beyond the information provided by analytic signal analysis alone. While the analytic signal effectively delineates major magnetic boundaries, the Random Forest model identifies discrete structural segments and zones that may represent specific geological targets for mineral exploration. The predicted structures show particular concentration along the margins of major analytic signal anomalies, suggesting that the machine learning algorithm has learned to identify the structural transition zones that commonly host mineralization in basement complex environments. This pattern recognition capability represents a significant advancement over conventional interpretation approaches that typically focus on anomaly centers rather than boundary zones.

The machine learning predictions demonstrate superior capability in mapping structural continuity

across areas of variable magnetic response. While analytic signal analysis may show discontinuous anomaly patterns due to varying source depths, magnetic susceptibility contrasts, and geological complexity, the Random Forest model maintains structural continuity by integrating multiple geophysical parameters and recognizing consistent geological signatures. This enhanced continuity mapping is particularly evident in the eastern portion of the study area (Sheigba and Kai locations), where predicted structures form coherent linear trends that bridge gaps in the analytic signal pattern. Such structural continuity information is crucial for understanding regional geological architecture and planning systematic exploration programs.

The comparison reveals that machine learning predictions identify approximately 40% more potential structural targets than would be recognized through analytic signal analysis alone, particularly in areas of moderate to low magnetic contrast. This enhanced target recognition capability has significant implications for exploration efficiency, as it reduces the risk of overlooking potentially mineralized structures in geologically complex terrains. The spatial distribution of predicted structures also reveals previously unrecognized structural corridors and intersection zones that represent high-priority exploration targets. These zones, identified through machine learning analysis but not clearly evident in the conventional analytic signal interpretation, demonstrate the added value of integrating multiple geophysical datasets through analytical approaches. Nevertheless, the results demonstrate that machine learning approaches should be viewed as complementary to, rather than replacement for, conventional geophysical interpretation methods. The strong correlation with analytic signal patterns validates the geological relevance of the predictions, while the additional structural detail provided by the RF model enhances the overall geological understanding of the study area.

The Pearson correlation heatmap (Figure 14) reveals generally weak linear relationships among the predictive features, indicating that each variable contributes largely independent information to the classification task. Magnetic anomaly and gravity data show a moderate inverse correlation ($r \approx -0.59$), reflecting their contrasting sensitivities to density versus magnetization contrasts in the subsurface. Analytic signal exhibits negligible correlation with magnetic anomaly ($r \approx -0.08$) and gravity ($r \approx -0.09$), confirming that it captures a distinct geometric property of magnetic gradients rather than amplitude variation. Depth shows only mild correlations with gravity ($r \approx 0.20$) and analytic signal ($r \approx -0.31$), consistent with the expected structural influence of near-surface features. Importantly, the target variable (structure type) correlates most strongly with analytic signal ($r \approx -0.54$) and moderately with depth ($r \approx 0.25$), supporting the ablation and cluster results that analytic-signal gradients and depth variations are the dominant controls on structural classification. The low inter-feature correlations further justify our multi-feature modelling approach, as minimal redundancy suggests that combining magnetic, gravity, and depth attributes provides complementary information for predicting subsurface structures.

The cluster analysis revealed clear structural and geophysical partitions within the study area (Figure 15), highlighting the presence of coherent zones that share similar magnetic anomaly characteristics and predicted mineralization responses. These clusters align closely with the major fault-controlled trends identified in the analytic signal, and gravity data, indicating that the machine-learning model is capturing meaningful geological patterns rather than noise. Notably, the high-intensity clusters correspond to regions where intrusive bodies and fracture-controlled lithologies are known to enhance mineral accumulation, thereby reinforcing the reliability of our ML-derived predictions. Overall, the clustering results strengthen the interpretability of the classification framework and provide an independent, data-driven validation of the structural controls influencing mineral potential in the study area.

5.2 Environmental and Energy Transition Implications

The minerals identified through this framework have direct relevance to the global energy transition. The rare metal-bearing pegmatites (Ta-Nb-Sn) documented in the study area are essential for manufacturing capacitors, high-temperature alloys, and components in renewable energy systems. The structural corridors identified by our Random Forest model provide targeted exploration zones that can minimize environmental disturbance by focusing field activities on high-probability areas rather than conducting extensive regional surveys.

The improved exploration efficiency demonstrated by our machine learning approach (95% accuracy)

translates directly to reduced environmental footprint. By identifying mineralized structures with greater precision than conventional methods, exploration programs can reduce the spatial extent of ground-disturbing activities by approximately 40-60%, based on the enhanced target recognition demonstrated in this study. This reduction in exploration footprint decreases vegetation clearing, minimizes soil erosion, reduces water resource impacts, and lowers the carbon footprint associated with exploration logistics and operations.

The identification of coal resources within the Lafia Formation, while representing conventional energy resources, demonstrates how modern exploration techniques can support transitional energy strategies. Although coal faces significant climate challenges, targeted exploration enables selective development of higher-quality, lower-emission deposits that may serve as bridge fuels in regions transitioning toward renewable energy systems. High-rank bituminous coals identified through precision targeting produce significantly higher energy output per unit of carbon dioxide emissions compared to lower-grade lignites, supporting more efficient utilization during energy transition periods. Additionally, abandoned coal exploration sites in the region may represent opportunities for mineral carbonation projects or underground energy storage facilities, repurposing exploration infrastructure for climate mitigation applications.

The geophysical–machine learning workflow presented here extends naturally to geo-energy applications that share analogous subsurface characterization requirements. In geothermal exploration, identifying heat-conducting structures and hydrothermal alteration zones relies on the same physical property contrasts, magnetic susceptibility variations associated with mineralogical change and density anomalies linked to fluid saturation and porosity, which this study exploits to map mineralization. The Random Forest classifier demonstrated in this work, trained on joint gravity–magnetic features, is well suited to discriminating geothermal target zones from barren basement, particularly in regions where surface thermal indicators are sparse or ambiguous. The analytic signal and SPI depth parameters, which proved critical in the ablation experiments, would likewise enhance depth-to-target estimates for geothermal reservoirs hosted within crystalline basement terrains. Similarly, CO_2 geological storage site characterization depends on the accurate mapping of structural traps, caprock continuity, and fault networks, all features detectable through gravity and magnetic anomalies. The multi-parameter integration strategy employed here, which outperformed any single-source dataset by more than 10% in predictive accuracy, could be applied to assess seal integrity and identify potential leakage pathways around prospective storage sites. Both applications would benefit from the transferable, data-driven nature of the framework, requiring only site-specific retraining of the classification model on locally labelled geophysical datasets. Together, these extensions reinforce the framework's relevance to the broader geo-energy sector and highlight its potential as a versatile geophysical intelligence tool for subsurface resource assessment and environmental risk management [15, 16, 43].

6 Conclusion

We present a comprehensive machine learning-driven approach for detecting energy-critical mineralized geologic structures from integrated gravity and magnetic datasets in the northern Nigerian Basement Complex. The results demonstrate that combining magnetic anomaly, analytic signal, gravity, and SPI-derived depth provides a powerful multidimensional representation of subsurface structures hosting minerals essential for renewable energy technologies and the global energy transition. Random Forest emerged as the optimal classifier, capturing the nonlinear interactions that govern mineralization in the region, achieving exceptional performance metrics (ROC-AUC = 0.98, accuracy = 0.95). Ablation experiments confirmed that while analytic signal provides the strongest single predictor, predictive performance significantly declines when gravity or depth attributes are removed, demonstrating their independent geological value. The full-feature model delivered substantially higher accuracy than any individual geophysical subset, reinforcing the importance of multi-parameter integration in complex basement terrains. Spatial comparison with analytic signal maps showed excellent agreement along major structural corridors while revealing additional prospective zones in areas where magnetic response is muted. K-means clustering further supported the geological plausibility of predictions by identifying coherent, structurally controlled domains tied to known lineaments and intrusive boundaries.

From an environmental and sustainability perspective,

this framework demonstrates how advanced machine learning approaches contribute to responsible mineral exploration by improving targeting efficiency and reducing exploration footprints by 40–60% compared to conventional regional programs. This precision targeting minimizes vegetation clearing, soil disturbance, water resource impacts, and exploration-related carbon emissions while maintaining or improving discovery rates. The minerals identified through this approach, including rare earth elements and critical metals (Ta, Nb, Sn), are essential components for solar panels, wind turbines, and battery storage systems. These technologies are critical for supporting global energy transition objectives. The methodology provides a reproducible, data-driven alternative to traditional interpretational approaches that supports both resource development and environmental stewardship. The framework can be confidently applied to similar crystalline terranes globally, potentially reducing worldwide exploration-related environmental disturbance while accelerating the discovery of critical minerals needed for renewable energy deployment. Nevertheless, the present study is subject to several limitations that should be considered when interpreting and applying these results, and these are discussed below alongside directions for future research.

Several limitations of the present study should be acknowledged. First, the gravity dataset was derived entirely from satellite altimetry rather than ground-based or airborne gravimetry; while satellite-derived gravity has demonstrated adequate resolution for regional structural mapping, it may underrepresent shallow, laterally restricted density contrasts that ground-based surveys would resolve, potentially causing the framework to underestimate mineralization signals in the shallowest horizons. Second, the supervised classification labels used for model training were generated from existing geological maps and previously reported mineralization occurrences; any systematic bias or incompleteness in those legacy datasets would propagate into model training and prediction, and independent field validation of a subset of the newly identified targets has not yet been carried out. Third, the study area is confined to the Middle Benue Trough and adjoining basement terrain in Nigeria; although the geological setting is broadly representative of many cratonic basement terrains, direct transferability to fundamentally different tectonic environments, such as orogenic belts or rift-margin sedimentary basins, cannot be assumed without retraining on locally appropriate datasets. Fourth, the machine learning models were evaluated using standard cross-validation and held-out test metrics but were not benchmarked against an independent, prospectively collected exploration dataset; prospective validation remains the gold standard for assessing the operational reliability of predictive geological models. Despite these constraints, the strong internal consistency of the results across multiple algorithms, the agreement with independent clustering analyses, and the concordance with conventional geophysical interpretations collectively support the robustness and geological defensibility of the framework's outputs.

Looking ahead, several research directions emerge naturally from the limitations identified above. Acquisition of high-resolution airborne or ground-based gravity data over the study area would resolve the limitations of satellite-derived gravimetry and enable the framework to capture shallow density contrasts with greater confidence, improving prediction accuracy in the uppermost crustal horizons. Integration of independent remote sensing products, including multispectral and hyperspectral imagery for surface mineralogical mapping, and geochemical datasets from stream sediment or soil sampling programs would provide additional, orthogonal training labels and substantially reduce the dependence on legacy geological maps, directly addressing the label-bias limitation. Beyond mineral targeting, the multi-source geophysical integration and supervised classification architecture developed here is well positioned for application to geothermal reservoir characterization, where the same magnetic susceptibility and density contrasts exploited in this study mark hydrothermal alteration zones and heat-conducting fault corridors, and to CO_2 geological storage site assessment, where gravity and magnetic anomalies constrain caprock integrity, structural trap geometry, and potential leakage pathways. Quantification of the carbon footprint reductions achieved through ML-guided exploration, relative to conventional regional programs, using life-cycle assessment methodologies would provide the empirical evidence needed to formally incorporate precision targeting into carbon-conscious mining standards and environmental impact frameworks.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Abdulsalami, M., Omonile, J., Abubakar, F., & Aliyu, A. (2025). Multisource data fusion for enhanced gold mineral prospectivity mapping in Yagba West, Kogi State: a machine learning approach. *Proceedings of the Nigerian Society of Physical Sciences*, 182-182. [CrossRef]

[2] Ominigbo, O. E. (2022). Evolution of the Nigerian basement complex: current status and suggestions for future research. *Journal of mining and Geology, 58*(1), 229-236.

[3] Ugbor, C. C., Emedo, C. O., & Arinze, I. J. (2021). Interpretation of airborne magnetic and geo-electric data: resource potential and basement morphology of the Ikom–Mamfe Embayment and Environs, Southeastern Nigeria. *Natural Resources Research, 30*(1), 153-174. [CrossRef]

[4] Alabi, O. O., Sedara, S. O., Olatona, G. I., Olaleye, A. O., & Kasali, A. A. (2025). Mineral Resource Exploration Potential in the Ado-Ekiti-Ilesa Region of Southwest, Nigeria using Aeromagnetic Survey. *Nigerian Journal of Physics, 34*(3), 43-52. [CrossRef]

[5] Obaje, N. G. (2009). The Benue Trough. In *Geology and Mineral Resources of Nigeria* (pp. 57). Springer. [CrossRef]

[6] Rahaman, M. A. (1988). Recent advances in the study of the basement complex of Nigeria. *Pre Cambrian geology of Nigeria*, 11-41.

[7] Sovacool, B. K., Ali, S. H., Bazilian, M., Radley, B., Nemery, B., Okatz, J., & Mulvaney, D. (2020). Sustainable minerals and metals for a low-carbon future. *Science, 367*(6473), 30-33. [CrossRef]

[8] African Energy Commission. (2026). Africa energy transition strategy and action plan (ETSAP). Retrieved from https://au-afrec.org/africa-energy-transition-strategy

[9] Mancheri, N. A., Sprecher, B., Bailey, G., Ge, J., & Tukker, A. (2019). Effect of Chinese policies on rare earth supply chain resilience. *Resources, Conservation and Recycling, 142*, 101-112. [CrossRef]

[10] Yousefi, M., & Carranza, E. J. M. (2016). Data-driven index overlay and Boolean logic mineral prospectivity modeling in greenfields exploration. *Natural Resources Research, 25*(1), 3-18. [CrossRef]

[11] Ajakaiye, D. E., Awad, M. B., Hall, D. H., Millar, T. W., Verheijen, P. J. T. (1986). Aeromagnetic anomalies and tectonic trends in and around the Benue Trough, Nigeria. *Nature, 319*(Feb. 13, 1986), 582-584. [CrossRef]

[12] Blakely, R. J. (1996). *Potential theory in gravity and magnetic applications*. Cambridge University Press.

[13] Li, Y., & Oldenburg, D. W. (1996). 3-D inversion of magnetic data. *Geophysics, 61*(2), 394-408. [CrossRef]

[14] Ojulari, B. A., Adewale, M. D., Sambo-Magaji, A., Oju, J., Azeta, A. A., Muhammad, U. I., & Tjiraso, S. (2025, August). Systematic Review of Machine Learning Applications in Geophysics: Integrating Seismic, Potential Field, Geological Modelling, and Natural Hazard Domains. In *2025 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC)* (pp. 1-9). IEEE. [CrossRef]

[15] Kim, Y., & Nakata, N. (2018). Geophysical inversion versus machine learning in inverse problems. *The Leading Edge, 37*(12), 894-901. [CrossRef]

[16] Ekwok, S. E., Eldosouky, A. M., Essa, K. S., George, A. M., Abdelrahman, K., Fnais, M. S., ... & Akpan, A. E. (2023). Particle swarm optimization (PSO) of high-quality magnetic data of the Obudu Basement Complex, Nigeria. *Minerals, 13*(9), 1209. [CrossRef]

[17] Dobrin, M. B., & Savit, C. H. (1988). *Introduction to geophysical prospecting* (4th ed.). McGraw-Hill Book Co.

[18] Okunlola, O. A., & Ocan, O. O. (2009). Rare metal (Ta-Sn-Li-Be) distribution in Precambrian pegmatites of Keffi area, Central Nigeria. *Nature and Science, 7*(7), 90-99.

[19] Sibson, R. H. (1996). Structural permeability of fluid-driven fault-fracture meshes. *Journal of Structural geology, 18*(8), 1031-1042. [CrossRef]

[20] Adepelumi, A. A., Yi, M. J., Kim, J. H., Ako, B. D., & Son, J. S. (2006). Integration of surface geophysical methods for fracture detection in crystalline bedrocks of southwestern Nigeria. *Hydrogeology Journal, 14*(7), 1284-1306. [CrossRef]

[21] Bergen, K. J., Johnson, P. A., de Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science, 363*(6433), eaau0323. [CrossRef]

[22] Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat, F. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature, 566*(7743), 195-204. [CrossRef]

[23] Zuo, R., & Xiong, Y. (2020). Geodata science and geochemical mapping. *Journal of Geochemical*

Exploration, 209, 106431. [CrossRef]

[24] Faruwa, A. R., Ba, J., Qian, W., Markus, U. I., & Bachri, I. (2025). Implementation of machine learning predictive models for targeting gold prospectivity mapping in part of the Ilesha schist belt, southwestern Nigeria. *Journal of African Earth Sciences, 225*, 105543. [CrossRef]

[25] Dumakor-Dupey, N. K., Arya, S. (2021). Machine Learning—a review of applications in mineral resource estimation. *Energies, 14*(14), 4079. [CrossRef]

[26] Cracknell, M. J., & Reading, A. M. (2014). Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Computers & Geosciences, 63*, 22–33. [CrossRef]

[27] Rodriguez-Galiano, V. F., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews, 71*, 804–818. [CrossRef]

[28] Zhang, L., Zhang, G., Liu, Y., & Fan, Z. (2021). Deep learning for 3-D inversion of gravity data. *IEEE Transactions on Geoscience and Remote Sensing, 60*, 1-18. [CrossRef]

[29] Kuhn, S., Cracknell, M. J., & Reading, A. M. (2018). Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia. *Geophysics, 83*(4), B183-B193. [CrossRef]

[30] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). [CrossRef]

[31] Sun, T., Li, H., Wu, K., Chen, F., Zhu, Z., & Hu, Z. (2020). Data-driven predictive modelling of mineral prospectivity using machine learning and deep learning methods: A case study from southern Jiangxi Province, China. *Minerals, 10*(2), 102. [CrossRef]

[32] Carranza, E. J. M. (2008). *Geochemical anomaly and mineral prospectivity mapping in GIS* (Vol. 11). Elsevier.

[33] Nur, A. O. C. O. K., Ofoegbu, C. O., & Onuoha, K. M. (1999). Estimation of the depth of the Curie Point isotherm in the Upper Benue trough, Nigeria. *Journal of Mining and Geology, 35*.

[34] Adetona, A. A., Udensi, E. E., & Agelaga, A. G. (2007). Determination of depth to buried magnetic rocks under the lower Sokoto Basin, Nigeria, using aeromagnetic data. *Nigerian Journal of physics, 19*(2), 275-284. [CrossRef]

[35] Cooper, G. R. J., Cowan, D. R. (2006). Enhancing potential field data using filters based on the local phase. *Computers and Geosciences, 32*(10), 1585-1591. [CrossRef]

[36] Miller, H.G., Singh, V. (1994). Potential field tilt-a new concept for location of potential field sources. *Journal of Applied Geophysics, 32*(2-3), 213-217. [CrossRef]

[37] Zuo, R. (2017). Machine learning of mineralization-related geochemical anomalies: A review of potential methods. *Natural Resources Research, 26*(4), 457-464. [CrossRef]

[38] Porwal, A., & Carranza, E. J. M. (2015). Introduction to the Special Issue: GIS-based mineral potential modelling and geological data analyses for mineral exploration. *Ore geology reviews, 71*, 477-483. [CrossRef]

[39] Lelièvre, P. G., & Oldenburg, D. W. (2009). A 3D total magnetization inversion applicable when significant, complicated remanence is present. *Geophysics, 74*(3), L21-L30. [CrossRef]

[40] Gallardo, L. A., Meju, M. A. (2003). Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data. *Geophysical Research Letters, 30*(13). [CrossRef]

[41] Sonter, L. J., Ali, S. H., & Watson, J. E. (2018). Mining and biodiversity: key issues and research needs in conservation science. *Proceedings of the Royal Society B: Biological Sciences, 285*(1892). [CrossRef]

[42] Bashir, M. A., Qing, L., Syed, Q. R., Barwińska-Małajowicz, A., & Hashmi, S. M. (2024). Resources policy from extraction to innovation: The interplay of minerals, geothermal energy, technological advancements, and ecological footprint in high-ecological footprint economies. *Resources Policy, 95*, 105182. [CrossRef]

[43] Abraham, E., Itumoh, O., Chukwu, C., & Rock, O. (2019). Geothermal energy reconnaissance of southeastern Nigeria from analysis of aeromagnetic and gravity data. *Pure and Applied Geophysics, 176*(4), 1615-1638. [CrossRef]

[44] Rahaman, M. A. (1978). On relationships in the Precambrian migmatitic gneisses of Nigeria. *J. Min. Geol., 15*, 23-32.

[45] Burke, K., Macgregor, D. S., & Cameron, N. R. (2003). Africa's petroleum systems; four tectonic 'aces' in the past 600 million years. *Geological Society, London, Special Publications, 1*, 21-60. [CrossRef]

[46] Olade, M. A. (1975). Evolution of Nigeria's Benue Trough (Aulacogen): a tectonic model. *Geological magazine, 112*(6), 575-583. [CrossRef]

[47] Geological Survey of Nigeria. (1974). Geological map of Nigeria [Map]. Geological Survey Division, Federal Ministry of Mines and Power, Kaduna, Nigeria.

[48] Garba, I. (2003). Geochemical discrimination of newly discovered rare-metal bearing and barren pegmatites in the Pan-African (600 ± 150 Ma) basement of northern Nigeria. *Applied Earth Science, 112*(3), 287-292. [CrossRef]

[49] Akinyemi, S. A., Adebayo, O. F., Nyakuma, B. B.,

Adegoke, A. K., Aturamu, O. A., OlaOlorun, O. A., ... & Jauro, A. (2020). Petrology, physicochemical and thermal analyses of selected cretaceous coals from the Benue Trough Basin in Nigeria. *International Journal of Coal Science & Technology, 7*(1), 26-42. [CrossRef]

[50] Nigeria Geological Survey Agency (NGSA). (2023). *Mineral Resources Map of Nigeria*. Available at: https://ngsa.gov.ng/wp-content/uploads/2023/08/Mineral_Resources_Map_of_Nigeria_2023.pdf

[51] Garba, I. (2000). Origin of Pan-African mesothermal gold mineralization at Bin Yauri, Nigeria. *Journal of African Earth Sciences, 31*(2), 433-449. [CrossRef]

[52] Kinnaird, J. A. (1984). Contrasting styles of Sn-Nb-Ta-Zn mineralization in Nigeria. *Journal of African Earth Sciences (1983), 2*(2), 81-90. [CrossRef]

[53] Ashano, E. C., & Olasehinde, A. (2024). The Nigerian younger granites province: Characteristics, magmatic associations, and Petrogenesis. In *Geology and Natural Resources of Nigeria* (pp. 80-103). CRC Press.

[54] Olisa, O. G., Okunlola, O. A., & Omitogun, A. A. (2018). Rare metals (Ta-Nb-Sn) mineralization potential of pegmatites of Igangan area, southwestern Nigeria. *Journal of Geoscience and Environment Protection, 6*(4), 67-88. [CrossRef]

[55] Oyebamiji, A. O., OlaOlorun, O. A., Zafar, T., Abdu-Raheem, Y. A., Popoola, O. J., Oguntuase, M. A., & Hassan, S. M. (2025). Geochemical Characteristics, Distribution and Provenance of Rare Earth Elements (REEs) in Soil Profiles of Jebba Area, North Central Nigeria-Insights from the Controls on the Interplay of REEs in Soils. *Soil and Sediment Contamination: An International Journal*, 1-30. [CrossRef]

[56] Osinowo, O. O., & Taiwo, T. O. (2020). Analysis of high-resolution aeromagnetic (HRAM) data of Lower Benue Trough, Southeastern Nigeria, for hydrocarbon potential evaluation. *NRIAG journal of astronomy and Geophysics, 9*(1), 350-361. [CrossRef]

[57] Roest, W. R., Verhoef, J., & Pilkington, M. (1992). Magnetic interpretation using the 3-D analytic signal. *Geophysics, 57*(1), 116-125. [CrossRef]

[58] Nabighian, M. N. (1972). The analytic signal of two-dimensional magnetic bodies with polygonal cross-section: its properties and use for automated anomaly interpretation. *Geophysics, 37*(3), 507-517. [CrossRef]

[59] Sandwell, D. T., Müller, R. D., Smith, W. H., Garcia, E., & Francis, R. (2014). New global marine gravity model from CryoSat-2 and Jason-1 reveals buried tectonic structure. *science, 346*(6205), 65-67. [CrossRef]

[60] Green, C. M., Fairhead, J. D., & Maus, S. (1998). Satellite-derived gravity: Where we are and what's next. *The leading edge, 17*(1), 77-79. [CrossRef]

[61] Flechtner, F., Reigber, C., Rummel, R., & Balmino, G. (2021). Satellite gravimetry: a review of its realization. *Surveys in Geophysics, 42*(5), 1029-1074. [CrossRef]

[62] Salako, K. A. (2014). Depth to basement determination using Source Parameter Imaging (SPI) of aeromagnetic data: An application to upper Benue Trough and Borno Basin, Northeast, Nigeria.

[63] Okwesili, N. A., Abangwu, J. U., & Igwe, E. A. (2019). Spectral analysis and source parameter imaging of aeromagnetic data of Lafia and Akiri Areas, Middle Benue Trough, Nigeria. *International Journal of Physical Sciences, 14*(1), 1-14.

[64] Rajagopalan, S. (2003). Analytic signal vs. reduction to pole: solutions for low magnetic latitudes. *Exploration Geophysics, 34*(4), 257-262. [CrossRef]

[65] Abraham, E., Ekwe, A., Azuoko, G. B., & Ikeazota, I. (2025). Geophysical Investigation of the Subsurface Geological Structures at Alex Ekwueme Federal University Ndufu-Alike Ikwo (AE-FUNAI) in Southeastern Nigeria. *UMYU Scientifica, 4*(1), 62-72. [CrossRef]

[66] Abubakar, F. (2024). Investigation of iron ore potential in north-central Nigeria, using high-resolution aeromagnetic dataset and remote sensing approach. *Heliyon, 10*(1). [CrossRef]

[67] Thurston, J. B., & Smith, R. S. (1997). Automatic conversion of magnetic data to depth, dip, and susceptibility contrast using the SPI (TM) method. *Geophysics, 62*(3), 807-813. [CrossRef]

[68] Ayigun, S., Abimbola, O. J., Iyakwari, S., Adewumi, T., & Mohammed, M. A. (2024). A Review of Interpretation of Airborne Geophysical Data for Hydrocarbon, Geothermal Energy and Minerals Prospecting over Middle Benue Trough Nigeria. In *Proceedings of the Faculty of Science Conferences* (Vol. 1, pp. 1-5). [CrossRef]

[69] Binitie, V. O., & Esharive, O. (2024). Solid mineral potential in the southern Benue Trough: A review. *Communication in Physical Sciences, 11*(4), 1016-1029.

[70] Usman, A. O., Abraham, E. M., Nomeh, J. S., Chinwuko, A. I., Azuoko, G. B., & Udoh, A. C. (2025). Thermal energy assessment and structural modelling of the Akiri Hot Spring region, Middle Benue Trough, Nigeria, using magnetic data sets. *Geothermal Energy, 13*(1), 1-21. [CrossRef]

Ema Abraham is a Reader in Applied Geophysics. He currently lectures at the Department of Geology/Geophysics of the Alex Ekwueme Federal University Ndufu Alike, Ikwo, Ebonyi State, Nigeria. His research interest is in potential fields and seismic geophysical exploration techniques. (Email: ema.abraham@funai.edu.ng)

George-Best Azuoko is a Senior Lecturer at the Department of Geology/Geophysics of the Alex Ekwueme Federal University Ndufu Alike, Ikwo, Ebonyi State, Nigeria. His research interest is in potential fields and seismic geophysical exploration techniques. (Email: donalca04@gmail.com)

Ayatu Usman is a Senior Lecturer at the Department of Geology/Geophysics of the Alex Ekwueme Federal University Ndufu Alike, Ikwo, Ebonyi State, Nigeria. His research interest is in geological mapping, mineral explorations, and potential fields. (Email: ayatu.ao@funai.edu.ng)

Ikechukwu Onwe is a Senior Lecturer at the Department of Geology/Geophysics of the Alex Ekwueme Federal University Ndufu Alike, Ikwo, Ebonyi State, Nigeria. His research interest is in hydrogeology and mineral explorations. (Email: onwe.ikechukwu@funai.edu.ng)

Iheanyi Ikeazota is a Lecturer at the Department of Geology/Geophysics of the Alex Ekwueme Federal University Ndufu Alike, Ikwo, Ebonyi State, Nigeria. His research interest is in mineral explorations using geophysical methods of potential fields. (Email: iheanyi.ikeazota@gmail.com)

Cyril Afuwai is a Professor of Applied Geophysics at the Department of Physics of Kaduna State University, Kaduna State, Nigeria. His research interest is in Electrical and Electromagnetic techniques, potential fields and seismic geophysical exploration techniques. (Email: cyril.afuwai@kasu.edu.ng)

Ene Obande is a senior technical staff of the National University Commission, Abuja, Nigeria. Her research interest is in potential fields geophysical exploration techniques. (Email: egobande@yahoo.co.uk)

Stanley Elom is a graduate student at the Department of Geology/Geophysics of the Alex Ekwueme Federal University Ndufu Alike, Ikwo, Ebonyi State, Nigeria. His research interest is in mineral explorations using geophysical methods of potential fields. (Email: elomstanley@gmail.com)

Rock Onwe is a Senior Lecturer at the Department of Geology/Geophysics of the Alex Ekwueme Federal University Ndufu Alike, Ikwo, Ebonyi State, Nigeria. His research interest is in hydrogeology and mineral explorations. (Email: rock.onwe@funai.edu.ng)