Bulletin of the Mineral Research and Exploration

http://bulletin.mta.gov.tr

Application of machine learning to hydrothermal system analysis: geochemical insights from the Bektakari-Bneli Khevi Ore Knot, Southern Georgia

Giorgi MINDIASHVILI[a*], David BLUASHVILI[b], Giorgi IOBIDZE[b], Tornike LIPARTIA[b], Nino JAFARIDZE[b] and Keti BENASHVILI[b]

[a] *Ivane Javakhishvili Tbilisi State University, Department of Geology, Tbilisi, Georgia.*
[b] *Georgian Technical University, Department of Applied Geology, Tbilisi Georgia.*

Research Article

Keywords:
Alteration Zones, Geochemical Analysis, HydrothermalSystems, Machine Learning, Principal Component Analysis (PCA).

ABSTRACT

This study integrates geochemical, statistical, and machine learning methods to investigate hydrothermal systems and mineralization processes within southern Georgia's Bektakari-Bnelikhevi ore knot. A total of 212 geochemical samples were analyzed, revealing key elemental associations such as V-Sc, Mo-W, and S-V, indicative of magmatic-hydrothermal activity and metasomatic alteration, including albitization and potassic enrichment. Principal Component Analysis (PCA) and DBSCAN clustering identified two dominant alteration regimes: Sulfide-rich mineralization and alkali metasomatism. Geochemical indices, Alteration Index (AI) and Chlorite-Carbonate-Pyrite Index (CCPI), effectively delineate alteration zones. AI values ranged from 45 to 95, while CCPI ranged from 30 to 85, with the highest mineralization potential concentrated in sericitic and Na-Ca zones. Feature importance analysis highlighted the Cu-Ag-Pb Index (32%) and Metallicity Factor (27%) as the strongest predictors of mineralized zones. Machine learning models achieved high precision in identifying epithermal and porphyry zones (Precision > 0.85), though recall remained low in transitional areas (Recall ~0.38), suggesting underrepresentation or overlapping features in these zones. This integrated approach offers a data-driven framework for targeting hydrothermal mineralization. The findings can inform exploration strategies by prioritizing geochemical signatures and improving zone classification in complex alteration systems.

Received Date: 23.12.2024
Accepted Date: 01.08.2025

1. Introduction

Hydrothermal systems play a central role in the formation and spatial distribution of mineral deposits, as they are closely associated with tectonic, magmatic, and fluid-driven processes (Burnham, 1979; Richards, 2003; Sillitoe, 2010). These systems contribute to diverse mineral assemblages and characteristic geochemical patterns that reflect their evolutionary history and physicochemical conditions (Gustafson and Hunt, 1975). Understanding hydrothermal mineralization requires a multidisciplinary approach that integrates geochemistry, mineralogy, and quantitative data analysis.

This study focuses on the Bektakari-Bnelikhevi ore knot in Southern Georgia, an area known for its complex geological structure and hydrothermal alteration. A total of 212 rock samples were collected and analyzed using Inductively Coupled

Citation Info: Mindiashvili,G., Bluashvili, D., Iobidze, G., Lipartia, T., Jafaridze, N., Benashvili, K. 2026. Application of machine learning to hydrothermal system analysis: geochemical insights from the Bektakari-Bneli Khevi Ore Knot, Southern Georgia. Bulletin of the Mineral Research and Exploration 179, 107-123. https://doi.org/10.19111/bulletinofmre.1768420

**Corresponding author: Giorgi MINDIASHVILI, giorgim1994@gmail.com*

Plasma-Atomic Emission Spectroscopy (ICP-AES). These data enable detailed characterization of the geochemical environment, helping to reveal the mechanisms behind hydrothermal mineralization and alteration processes (Meinert et al., 2005).

One of the key objectives of this research is to investigate vertically controlled mineralization patterns that reflect changes in pressure, temperature, and fluid-rock interaction with depth (Lowell and Guilbert, 1970). Particular attention is given to transition zones -geological environments where metasomatic and tectonic processes interact- since these areas often host complex alteration and mineralization trends (Richards, 2003).

Hydrothermal alteration zones are inherently heterogeneous, both spatially and geochemically, often resulting in irregular metal distributions (Pirajno, 2009). These complexities present analytical challenges, particularly when interpreting element associations and identifying the boundaries of mineralized zones. While previous studies have utilized geostatistical approaches to address such issues (Goovaerts, 1997), the present research applies statistical and machine learning methods as core analytical tools. Techniques such as Principal Component Analysis (PCA), clustering algorithms (e.g., DBSCAN), and feature importance analysis provide a data-driven framework for extracting meaningful patterns from complex geochemical datasets.

By integrating geochemical indices, such as the Alteration Index (AI) and Chlorite-Carbonate-Pyrite Index (CCPI), with machine learning classification models, this study aims to define alteration zones more precisely, interpret mineralization trends more robustly, and contribute to the development of improved mineral exploration strategies.

The study area, the Bektakari-Bnelikhevi ore knot, is located in southern Georgia, near the borders with Armenia and Türkiye (Figure 1). This region forms part of the Lesser Caucasus metallogenic belt,

Figure 1- Location of study area (Eurasian geopolitics.com, Walker, 2015).

known for its complex tectono-magmatic evolution and hydrothermal mineral systems. The yellow star on the map indicates the central position of the ore knot. Geodynamically, this segment lies within the convergent boundary of the Eurasian and Arabian plates, shaped by a long history of subduction, obduction, and continental collision processes (Sosson et al., 2010). The metallogeny of this zone is characterized by arc-related volcanic and subvolcanic rocks, hosting diverse base and precious metal mineralization (Mederer et al., 2014).

1.1. Geological Background

The Bektakari-Bneli Khevi ore knot is situated within the southern segment of the Lesser Caucasus metallogenic belt and is predominantly composed of Upper Cretaceous volcanic and volcano-sedimentary formations. According to the regional geological map (Figure 2), the study area is primarily underlain by rocks of the Didgverdi Suite (K2), which consists of Turonian-age rhyolites and rhyodacites, as well as basaltic-diabasic, microgabbroic intrusions. These igneous rocks are spatially associated with zones of hydrothermal alteration and serve as key host units for ore mineralization.

Surrounding the main study area are the Gasandami (K5) and Mashavera (K3-K5) suites, composed predominantly of Campanian to Senonian volcanoclastic and sedimentary rocks. These lithologies offer significant geochemical and structural contrasts and provide favorable horizons for hydrothermal fluid interaction and mineral deposition.

Fault structures are well-defined and appear to exert first-order control on the hydrothermal system. Both mapped and inferred faults traverse the alteration zones and strongly influence the localization of mineralized bodies. The ore zones, as highlighted on the structural map (in red and magenta), display clear alignment with these fault systems and lie near intrusive rhyolitic bodies. This spatial association implies a strong genetic link between magmatic intrusions and ore deposition processes.

The area hosts eight drillholes, shown as red circles on the map, strategically distributed across the Bektakari and Bneli Khevi subzones. Core samples from these drill holes formed the basis for this study: a total of 212 samples were collected and analyzed using Inductively Coupled Plasma-Optical Emission Spectroscopy (ICP-OES), enabling a high-resolution geochemical assessment of alteration and mineralization processes.

The tectono-magmatic and lithological configuration of this region reflects a typical volcanic arc setting. Successive pulses of arc magmatism, extensional faulting, and hydrothermal fluid migration have together created a favorable environment for polymetallic hydrothermal systems and associated mineral deposits (Adamia et al., 2017).

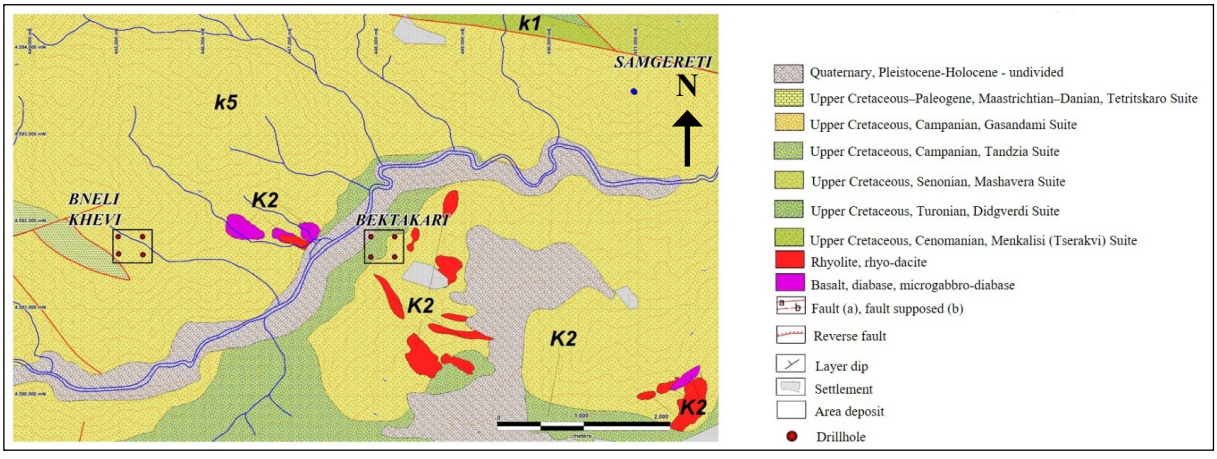

Figure 2- Geological map of Bektakari-Bnelikhevi ore knot (Adamia et al., 2017).

2. Method of Study

This study adopted an integrated methodological framework combining geochemical, statistical, and machine learning techniques to analyze hydrothermal alteration and mineralization processes in the Bektakari-Bnelikhevi ore knot, southern Georgia (Rollinson, 1993; Goovaerts, 1997; Jolliffe and Cadima, 2016). A total of 212 geochemical samples were collected from eight drill holes. These core samples were carefully prepared and subsequently analyzed using Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES). The elemental concentrations were determined using Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES) (Zhao et al., 1994).

2.1. Data Preprocessing

Before analysis, the geochemical dataset underwent several preprocessing steps. Missing values were imputed using mean substitution, outliers were identified and removed using the interquartile range (IQR) method, and all features were standardized using z-score normalization to ensure comparability. These steps ensured the reliability of statistical outputs and the stability of machine learning models (Jolliffe and Cadima, 2016).

2.2. Correlation and Dimensionality Reduction

A Pearson correlation matrix was calculated and visualized as a heatmap to explore elemental associations (Rollinson, 1993). This analysis revealed both positive and negative linear correlations, aiding in the identification of geochemical links and alteration trends. To reduce dimensionality and reveal key patterns, Principal Component Analysis (PCA) was applied. PCA transforms correlated variables into a set of uncorrelated principal components, emphasizing the main sources of variance. A biplot of the first two components allowed for the visualization of sample clusters and element loadings.

2.3. Clustering and Pattern Recognition

Unsupervised learning techniques were applied to identify natural groupings within the data. Hierarchical clustering using Ward's linkage method revealed broad elemental clusters (Franklin et al., 2005). Additionally, two machine learning clustering algorithms were employed: K-means clustering ($k = 2$) distinguished broad alteration regimes. DBSCAN (Density-Based Spatial Clustering of Applications with Noise), using eps = 1.5 and min_samples = 5, detected high-density clusters without requiring a predefined number of groups. Cluster results were projected into PCA space and visualized through scatter plots, allowing interpretation of compositional variability and geochemical heterogeneity.

2.4. Feature Selection and Classification

To determine the elements most strongly associated with mineralized zones, a feature importance analysis was conducted using a Random Forest classifier (Breiman, 2001). This method ranks variables based on their contribution to model performance. The Cu-Ag-Pb index and a composite Metallicity Factor emerged as dominant predictors. To evaluate the distribution of key metals (e.g., Au, Cu, Pb, Zn, Fe), box plots were generated to visualize central tendencies and detect geochemical anomalies (Boyle, 1979).

2.5. Geochemical Indices and Zonation

Two established geochemical indices were used to characterize alteration intensity: Alteration Index (AI): reflects sericitization and potassic alteration, based on mass changes in Na, Ca, and K (Ishikawa et al., 1976). Chlorite-Carbonate-Pyrite Index (CCPI): measures chloritic and carbonate alteration using Mg, Fe, and Na contents (Large et al., 2001). Data were plotted in AI-CCPI space and grouped into four alteration zones: Sericitic, intermediate, Na-Ca, and chloritic. These zones were classified using supervised machine learning, and classification performance was evaluated with precision, recall, and F1-score metrics (Powers, 2011).

2.6. Software and Optimization

All analyses were conducted using Python (version 3.10). Core libraries included: Pandas for data handling; Scikit-learn for machine learning and preprocessing; Matplotlib and Seaborn for visualization. Model optimization was performed using Grid Search Cross-Validation, which systematically tests multiple

hyperparameter combinations to maximize model performance (Pedregosa et al., 2011).

This methodology enables a comprehensive, multi-scale interpretation of hydrothermal geochemical systems. By integrating statistical analysis, unsupervised clustering, supervised learning, and geochemical indices, the study reveals robust spatial patterns and predictive indicators relevant to mineral exploration (Hedenquist and Lowenstern, 1994; Heinrich, 2007).

Workflow diagram (Figure 3) summarizing the integrated methodology used for hydrothermal mineralization analysis, combining geochemical preprocessing, dimensionality reduction, clustering, feature importance assessment, and alteration zonation through geochemical indices.

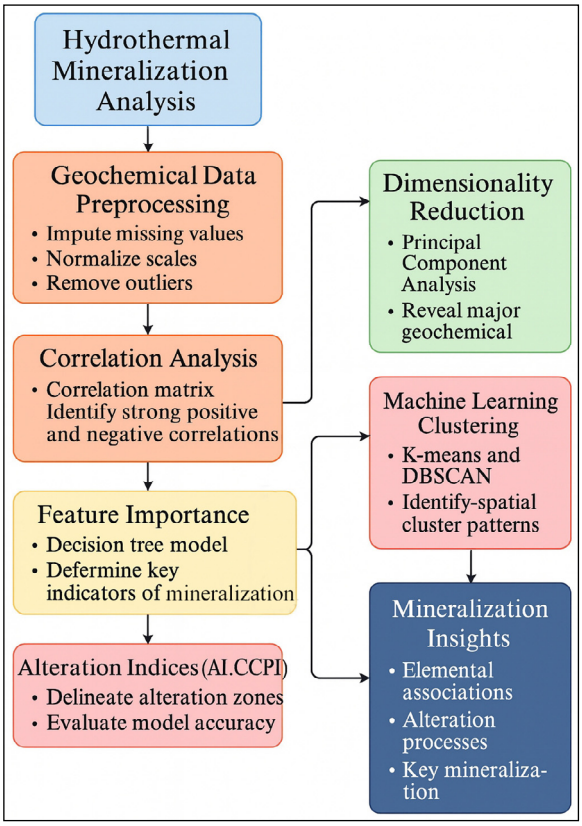

Figure 3- Schematic workflow of the analytical framework used to interpret hydrothermal mineralization in the Bektakari-Bneli Khevi ore knot. The process integrates geochemical data preprocessing (e.g., imputation, normalization, outlier removal), statistical analyses.

3. Results

3.1. Correlation Matrix Analysis

V and Sc (0.95): These elements are often linked to magmatic processes and siderophile elements (Sun and McDonough, 1989). Their strong correlation suggests a pronounced magmatic influence, potentially indicating magmatic mineralization or the presence of phosphate- and vanadium-rich rocks (Figure 4).

A strong correlation between molybdenum (Mo) and tungsten (W) (0.93) suggests the presence of hightemperature hydrothermal systems (Seedorff et al., 2005). These elements are frequently associated with porphyry deposits or skarn-type systems, where molybdenite (MoS_2) and wolframite [(Fe, Mn) WO_4] commonly occur.

S and V (0.93): The strong correlation between sulfur (S) and vanadium (V) may be attributed to the presence of sulfide minerals, including pyrite (FeS_2) or vanadium-bearing sulfides (Huston et al., 1995). This relationship highlights the intensity of sulfide mineralization processes within hydrothermal systems.

W and Ni (0.92): The strong association between tungsten (W) and nickel (Ni) suggests the presence of high-temperature hydrothermal processes. Nickel commonly forms through metamorphic, metasomatic, and hydrothermal activity, and its correlation with tungsten often indicates skarn systems or subduction-related environments (Meinert, 1992).

V and Co (0.92): The strong correlation between vanadium (V) and cobalt (Co) suggests the involvement of hydrothermal systems enriched in these metals (Groves et al., 1998). They commonly coexist with sulfide or oxide minerals, indicating pronounced mineralization processes.

K and Na (-0.77): This negative correlation suggests potassium metasomatic processes linked to albitization (the formation of sodium-rich feldspar) (Pirajno, 2009). Such processes often occur during intense hydrothermal alteration.

K and Mn (-0.66): This correlation may indicate the formation of K-rich phases resulting from metasomatic alteration (Rollinson, 1993), such as

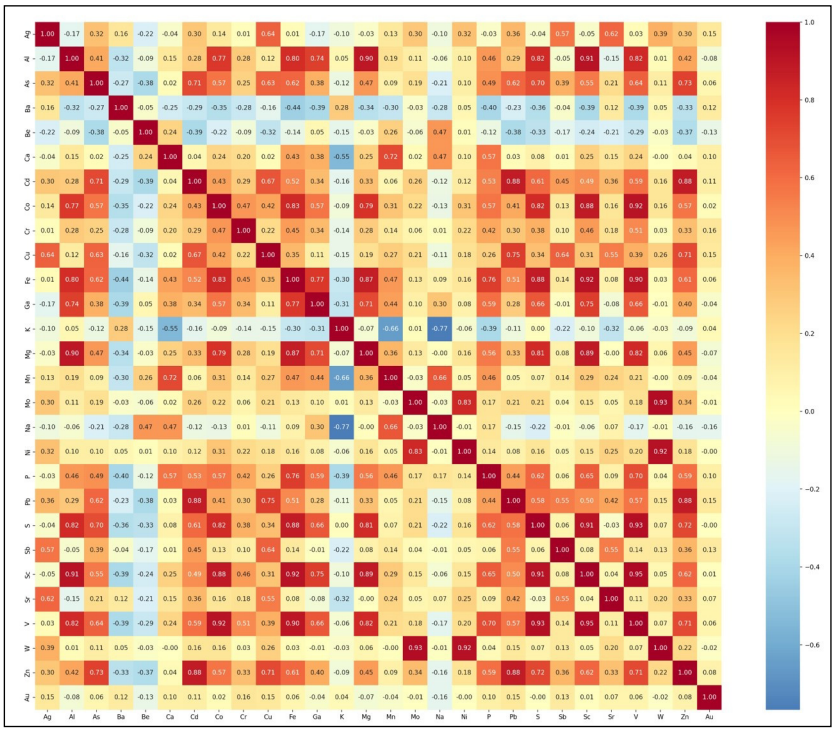

Figure 4- Correlation matrix illustrating the linear relationships among major and trace elements within the Bektakari-Bneli Khevi ore knot dataset. Strong positive correlations (e.g., V-Sc, Mo-W, S-V) suggest coupled geochemical behavior associated with high-temperature magmatic-hydrothermal systems. Negative correlations (e.g., K-Na, K-Mn) reflect metasomatic processes, such as sodic and potassic alteration. These relationships aid in the interpretation of mineralization style, elemental paragenesis, and fluid evolution.

orthoclase or biotite, which typically occur in high-temperature environments.

K and Ca (-0.55): This negative relationship suggests the formation of biotite and potassium feldspar (Barton et al., 1991), mineral phases commonly associated with metasomatic alteration processes.

Ba and Fe (-0.44): The negative correlation between barium (Ba) and iron (Fe) suggests their release as barite ($BaSO_4$) and pyrite (FeS_2), which is characteristic of sulfide- and barium-rich hydrothermal systems.

Ba and P (-0.40): This correlation may suggest the formation of barium phosphates, which can crystallize under the influence of hydrothermal solutions.

The V-Sc-Co-S association indicates metal-rich sulfide mineralization environments such as pyrite (FeS_2) or chalcopyrite ($CuFeS_2$), characteristic of high-temperature hydrothermal systems (Groves et al., 1998). Similarly, Mo-W-Ni co-associations are indicative of high-temperature porphyry or skarn-type mineralization (Evans, 1993). Negative correlations between K and Na, Mn, and Ca suggest metasomatic alterations, including the formation of orthoclase and biotite under chemically active hydrothermal regimes (Heinrich, 2007).

These trends are consistent with those observed in the correlation matrix, which revealed strong positive and negative relationships among key elements. These associations support interpretations of both high-temperature magmatic inputs and subsequent metasomatic overprinting (Sillitoe, 1993; Hedenquist and Lowenstern, 1994).

3.2. Principal Component and Clustering Analysis

The PCA biplot (Figure 5) illustrates variance among geochemical signatures. PC1 accounts for

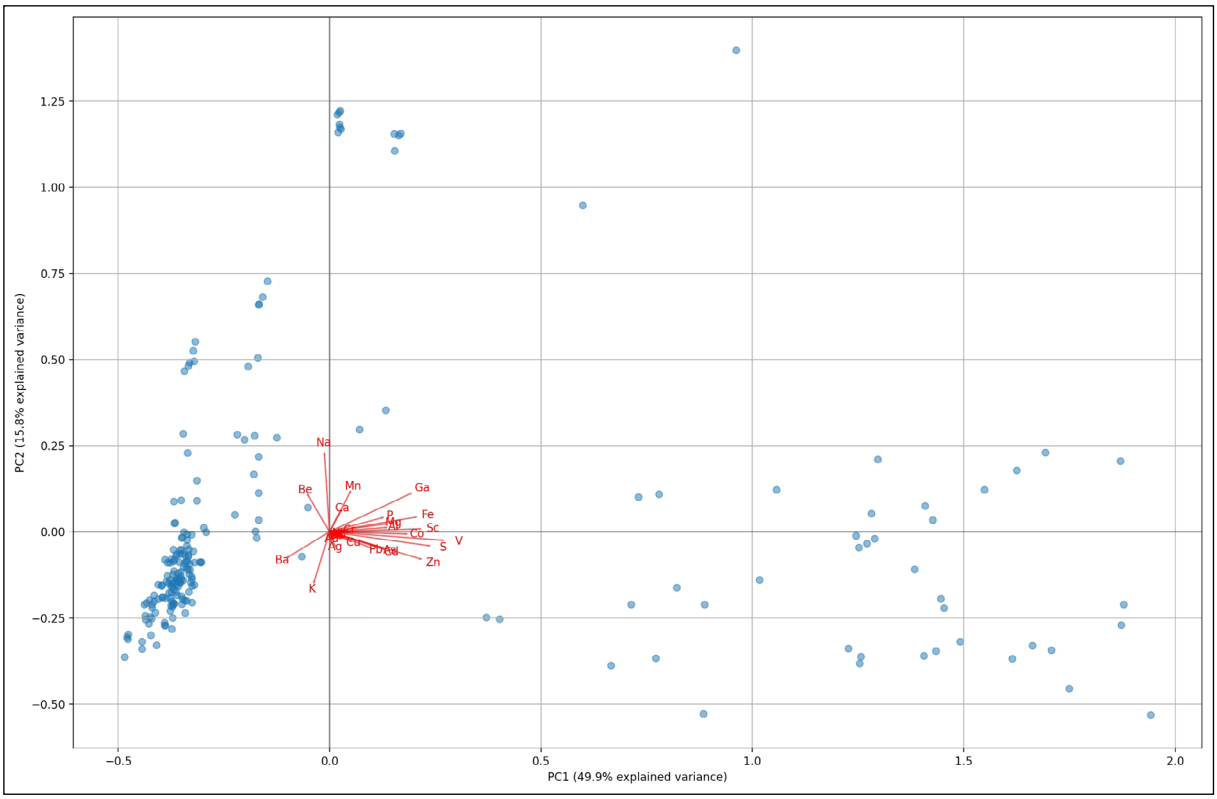

Figure 5- Principal Component Analysis (PCA) biplot illustrating the geochemical variability in the studied samples. PC1 (49.9% of explained variance) captures the major variance associated with high-temperature mineralization processes, primarily driven by Zn, S, Sc, and V. PC2 (15.8%) reflects metasomatic alterations, with key contributions from Na, K, and Ba. Red vectors represent the loadings of individual elements, indicating their influence on each component and aiding in the interpretation of underlying mineralization and alteration trends.

49.9% of the total variance and is dominated by V, Zn, S, and Sc—elements associated with high-temperature sulfide mineralization (Jolliffe and Cadima, 2016). The Zn-S covariance reflects sphalerite formation, a typical magmatic sulfide mineral (Cooke et al., 2000). PC2 explains 15.8% of the variance and reflects metasomatic changes, with Na, K, and Ba contributing strongly, suggesting feldspar alteration and barite precipitation (Kesler, 1997).

Hierarchical clustering (Figure 6) supports these interpretations and identifies three major geochemical groups:

Sulfide-related group (V, S, Zn, Sc): High-temperature mineralization (Huston et al., 1995; Franklin et al., 2005).

Alkaline/metasomatic group (Na, K, Ba): Potassic and sodic alteration (Heinrich, 2007).

Polymetallic group (Ag, Cu, Pb, Au): Reflects late-stage mineralization processes (Sillitoe, 1993; Hedenquist and Lowenstern, 1994).

These groups represent a mineralization evolution sequence from magmatic sulfides through metasomatic alteration to polymetallic enrichment (Nesbitt and Young, 1982; Taylor and McLennan, 1985;).

3.3. Metal Concentration by Gold Zones

As shown in Figure 7, elemental concentrations vary significantly across Au zones: High-Au zones are enriched in Zn, Cu, Pb, Ag, and Fe, indicating sulfide and polymetallic mineralization; Medium-Au zones: Elevated Mo and W concentrations reflect intermediate-stage mineralization (Evans, 1993); Low-Au zones: Exhibit background-level concentrations across all elements.

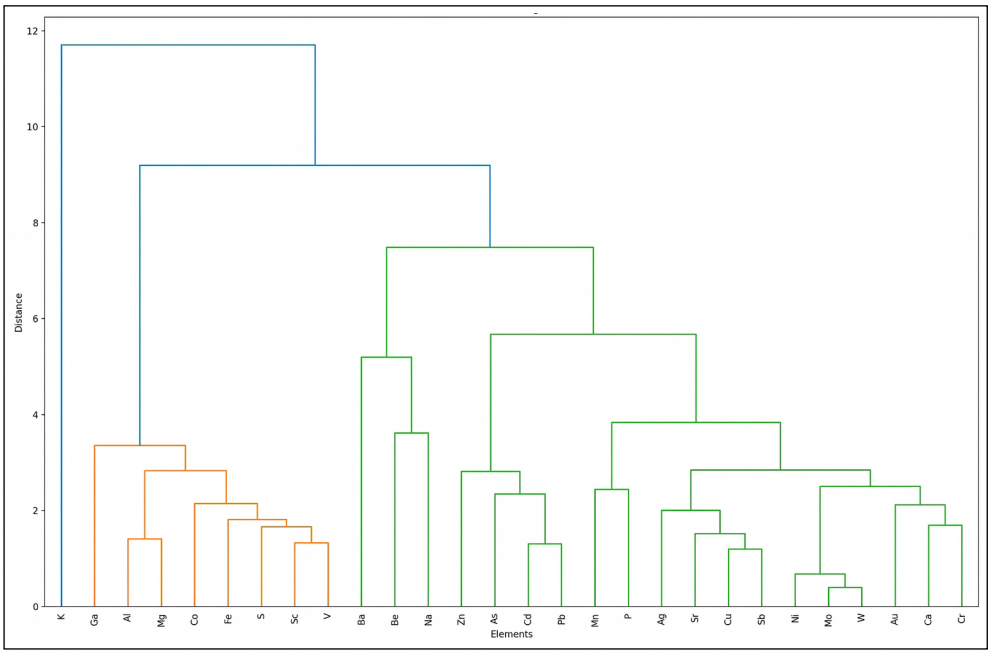

Figure 6- Hierarchical clustering dendrogram illustrating the relationships among geochemical elements based on compositional similarity. The clustering was performed using Ward's linkage method and Euclidean distance metric. Three major elemental groupings are evident: (1) a sulfide-associated cluster (e.g., V, S, Sc, Zn), (2) an alkali and metasomatic alteration cluster (e.g., Na, K, Ba, Mg), and (3) a polymetallic cluster representing late-stage mineralization (e.g., Cu, Ag, Pb, Au, Mo, W). These clusters correspond to different mineralization and alteration processes within the hydrothermal system.

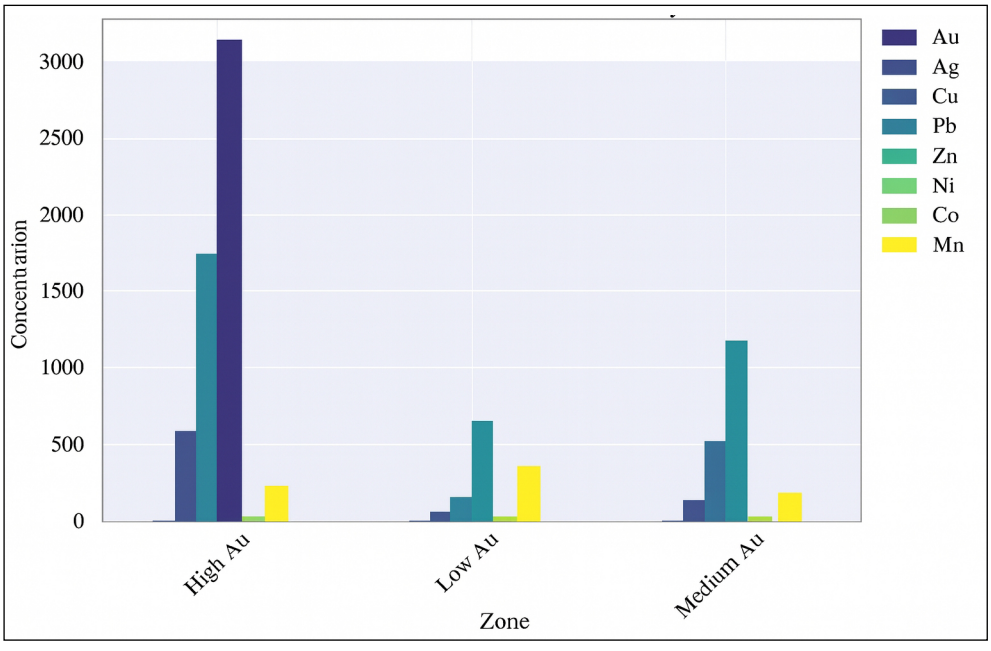

Figure 7- Mean concentrations of selected metals (Au, Ag, Cu, Pb, Zn, Ni, Co, Mn) across zones of varying gold content (High Au, Medium Au, Low Au). The plot illustrates elevated concentrations of Zn, Pb, and Cu in the High Au zone, indicating their association with sulfide and polymetallic mineralization phases. Medium Au zones show enrichment in Zn and Cu, while Low Au zones reflect background levels of metal concentration. This zonal differentiation provides insight into the temporal and spatial evolution of mineralizing fluids within the hydrothermal system.

Boxplots in Figure 8 confirm that Au, Cu, Pb, and Zn are concentrated in sericitic and intermediate zones, while Mo shows moderate enrichment in intermediate zones.

3.4. Alteration Zonation via AI and CCPI

AI vs. CCPI plotting (Figure 9) distinguishes four alteration types: Sericitic (high AI): Corresponds to high Au and base metals; Chloritic: Represents distal, Mg-rich alteration; Na-Ca: Early-stage, sodic alteration with low mineralization; Intermediate: Transitional zones with moderate enrichment.

These indices effectively zoned mineralization environments and aligned with metal distribution trends (Pearce et al., 1984).

3.5. Vertical and Lateral Controls

Structural domain control (Figure 10) shows: Deep porphyry zones: High vertical control (log Cu/Mo); Epithermal zones: High lateral control (log K/Na); Transition zones: Exhibit mixed structural characteristics.

These controls reflect tectono-magmatic evolution and fluid movement paths (Wilkinson, 2001).

3.6. Feature Importance and Classification Metrics

Feature importance results (Figure 11) highlight: Cu-Ag-Pb Index (30.8%) and Metallicity Factor (23.7%) as top predictors; Ag (23.1%) and Cu (22.2%) are strong standalone predictors; Pb (0.3%) contributes marginally but synergizes in combination.

Classification metrics (Figure 12) show: High performance in Sericitic and Na-Ca zones (Precision, Recall ~1.0); Lower recall in the Transitional zone (~0.38), despite moderate precision; Poor performance in Background zone (F1 ~0.32).

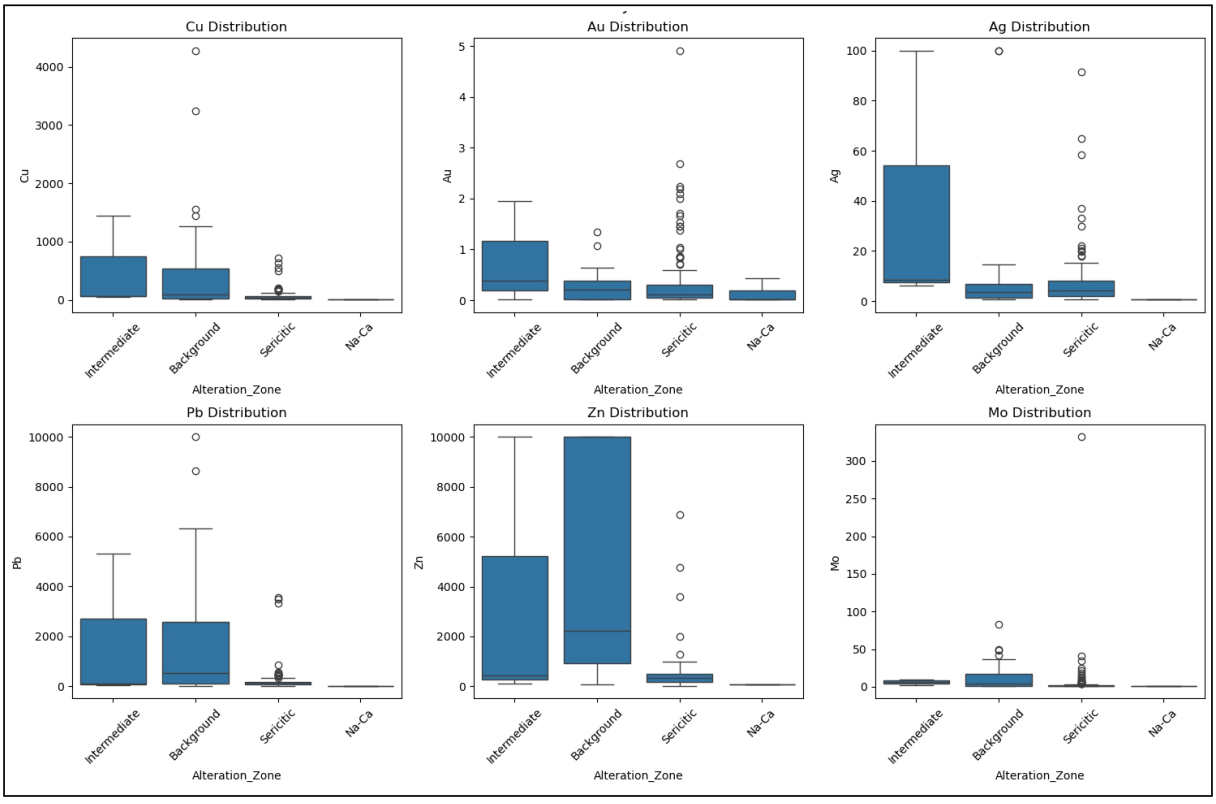

Figure 8- Boxplot distributions of key metals (Cu, Au, Ag, Pb, Zn, Mo) across alteration zones (Intermediate, Background, Sericitic, Na-Ca). The plots demonstrate significant variation in elemental concentrations between alteration environments. Notably, sericitic and intermediate zones are enriched in Au, Cu, Pb, and Zn, supporting their association with hydrothermal fluid pathways and mineralized zones. Mo shows moderate elevation in intermediate and background zones, while Na-Ca zones display consistently low metal concentrations, characteristic of early-stage or weakly mineralized environments. These trends reinforce the utility of alteration zoning as a predictive tool for mineral exploration.

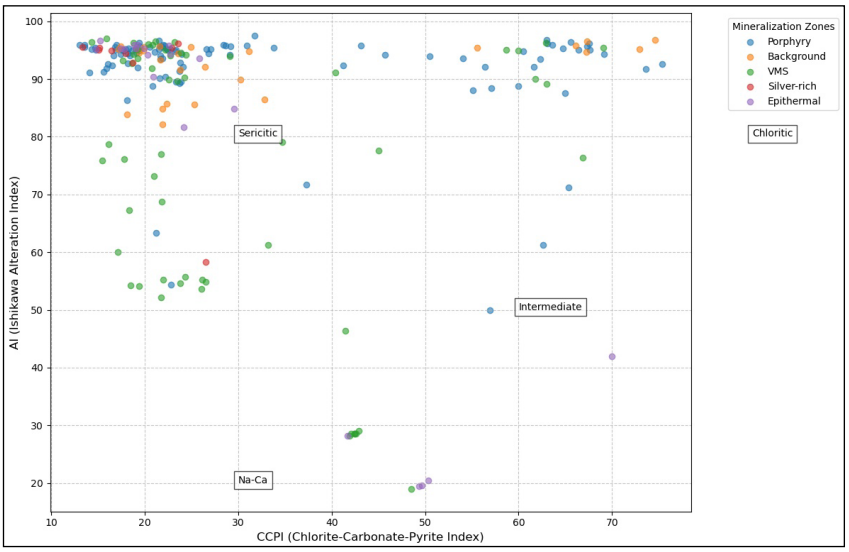

Figure 9- Alteration discrimination diagram based on Ishikawa Alteration Index (AI) and Chlorite-Carbonate-Pyrite Index (CCPI). The plot delineates four main alteration zones: Sericitic, Intermediate, Na-Ca, and Chloritic. Each point represents a sample colored by its associated mineralization type (Porphyry, VMS, Silver-rich, Epithermal, Background). Sericitic zones cluster at high AI and moderate CCPI values, while Na-Ca zones occupy the low end of both indices, consistent with early sodic alteration. Intermediate and chloritic zones span transitional index values. This spatial geochemical separation highlights the effectiveness of alteration indices for vectoring towards mineralized systems.

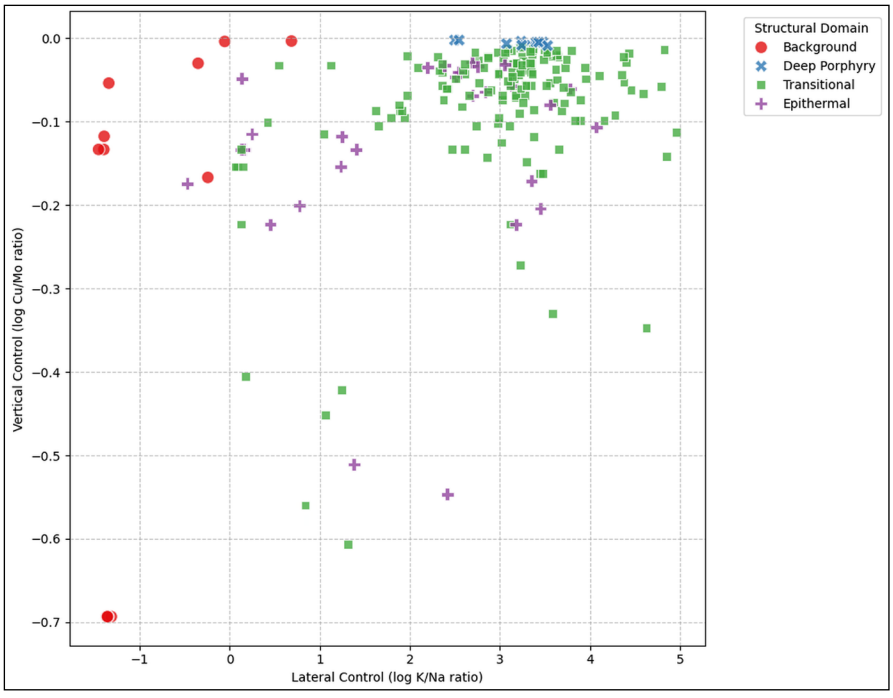

Figure 10- Plot of vertical versus lateral structural control across mineralization zones. The x-axis shows lateral control using the log-transformed K/Na ratio, while the y-axis displays vertical control via the log-transformed Cu/Mo ratio. Four structural domains are identified: Background (red circles), Deep Porphyry (blue crosses), Transitional (green squares), and Epithermal (purple pluses). Epithermal zones exhibit high lateral control, Deep Porphyry zones show higher vertical control, and Transitional zones occupy intermediate positions. This plot illustrates how elemental ratios can reveal spatial fluid migration pathways and structural influences on mineral deposition.

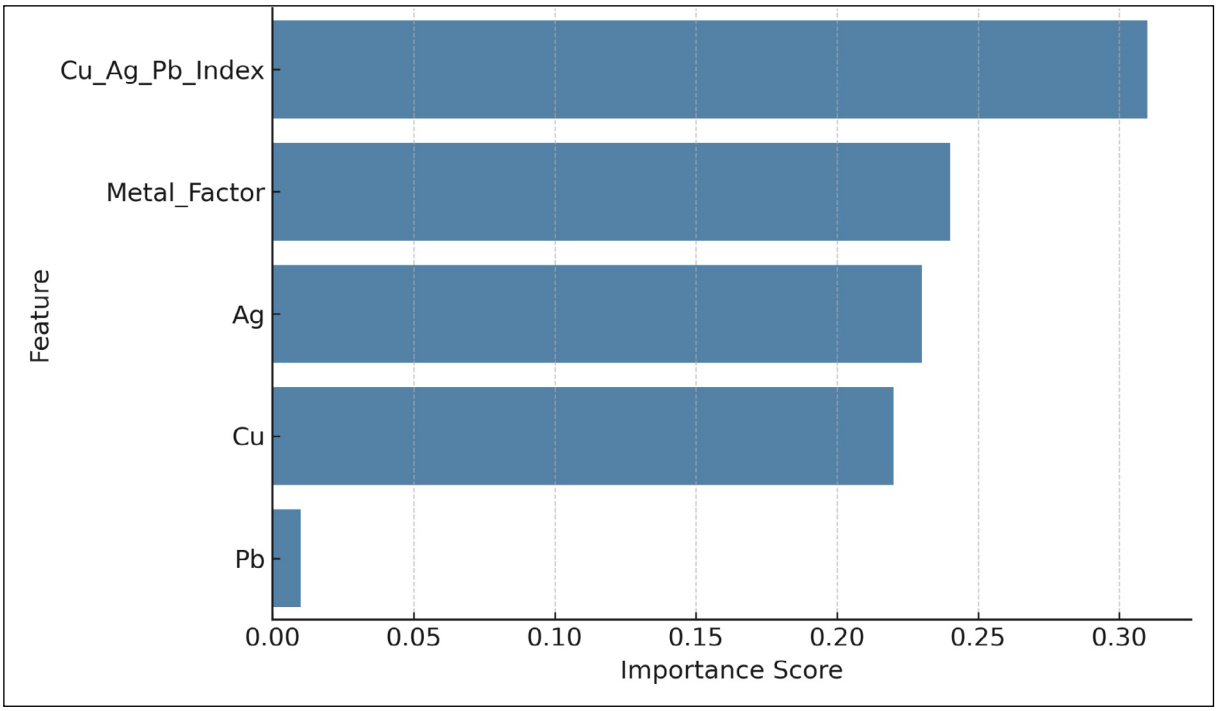

Figure 11- Feature importance plot from the Random Forest classification model, illustrating the relative contribution of selected geochemical indicators to mineralization zone prediction. The Cu-Ag-Pb Index emerged as the most influential predictor (importance score ~0.31), followed by the composite Metallicity Factor (~0.24), and individual contributions from Ag (~0.23), Cu (~0.22), and Pb (~0.003). These features highlight the strong predictive power of multi-element indices and the geochemical relevance of copper and silver in the studied hydrothermal system.

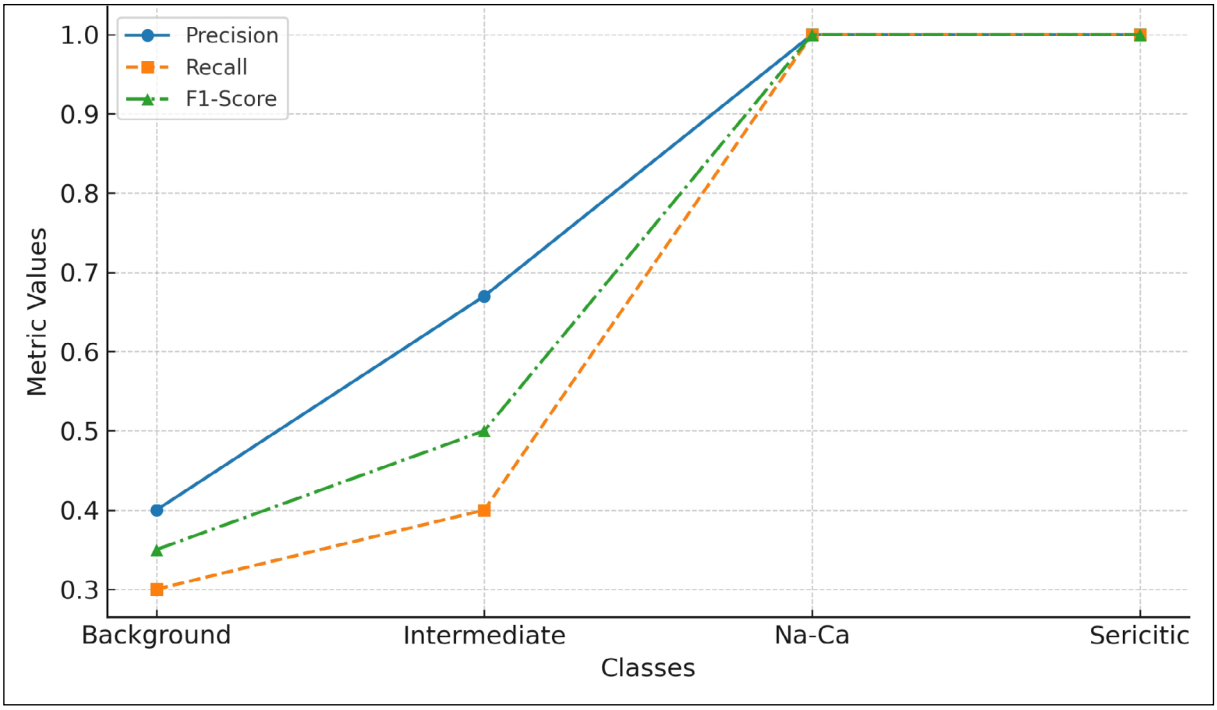

Figure 12- Classification metrics (Precision, Recall, F1-Score) for each alteration zone class predicted by the Random Forest model. The Sericitic and Na-Ca zones exhibit near-perfect classification performance (metrics ~1.0), indicating strong geochemical distinctiveness. In contrast, the Intermediate zone shows moderate classification scores (e.g., Recall ~0.38), and the Background zone performs poorly across all metrics (F1-Score ~0.32), reflecting overlapping or less distinctive geochemical signatures.

These results validate the utility of combined geochemical indices for classification, while also indicating areas where model refinement is needed (Powers, 2011; Saito and Rehmsmeier, 2015).

3.7. Statistical Summary of Geochemical Data

Descriptive statistics (Table 1) reveal significant variation among geochemical parameters, supporting the geochemical complexity of the hydrothermal system. Key elements such as Ag, As, Cu, Pb, Zn, and Mo display high standard deviations and coefficient of variation (CV%), indicating strong heterogeneity in mineralization and fluid chemistry (Boyle, 1979; Rollinson, 1993).

Silver (Ag): Mean = 7.68 ppm; CV% = 192% -Extreme variation in Ag values reflects late-stage hydrothermal overprinting, consistent with polymetallic and epithermal enrichment (Boyle, 1979).

Table 1- Descriptive statistics of major and trace elements in the Bektakari-Bneli Khevi ore knot based on 212 core samples analyzed by ICP-AES. The table includes mean, median, minimum, maximum, standard deviation, and coefficient of variation (CV%) for each element. High CV values (e.g., Ag, Cu, Pb, Zn, Mo) reflect strong geochemical heterogeneity associated with hydrothermal mineralization.

	Mean	Median	Minimum	Maximum	Standard Deviation	CV (%)
Ag	7.682547	3.6	0.5	100	14.76939	192.246
Al	7.04783	6.42	4.65	16.55	1.935072	27.45628
As	115.7642	55	5	1005	172.7496	149.2255
Ba	2164.151	1735	20	9040	1907.457	88.1388
Be	0.781132	0.8	0.5	1.6	0.200999	25.73176
Ca	0.202311	0.1	0.04	4.2	0.405367	200.3678
Cd	10.12689	0.5	0.5	149	27.44568	271.0179
Co	6.04717	1	1	52	10.51551	173.8914
Cr	6.853774	1	1	193	16.4807	240.4617
Cu	153.8208	34	1	4280	438.2335	284.8988
Fe	1.59184	0.79	0.25	7.88	1.640036	103.0277
Ga	13.06604	10	10	30	5.108805	39.09988
K	3.966557	4.13	0.74	5.87	1.021199	25.74524
Mg	0.550802	0.35	0.07	3.55	0.548518	99.58533
Mn	123.2358	71.5	22	970	148.9263	120.8466
Mo	6.688679	1	1	332	24.63126	368.2529
Na	0.636132	0.19	0.03	4.03	0.978876	153.8794
Ni	5.311321	2	1	164	13.85015	260.7667
P	146.4151	60	30	1240	200.8451	137.1751
Pb	611.7217	121.5	4	10000	1450.651	237.1422
S	0.949717	0.03	0.01	7.88	1.954464	205.7943
Sb	12.04717	6	5	551	38.65362	320.8523
Sc	9.537736	4	2	52	11.21371	117.5721
Sr	77.21226	45	6	1255	129.8466	168.1684
V	48.13679	9	1	302	83.14189	172.72
W	18.91509	10	10	660	48.33976	255.5618
Zn	1431.882	425.5	14	10000	2693.461	188.1063
Au	0.313538	0.12	0.01	4.91	0.554427	176.8293
AI	85.01677	93.54228	17.15135	97.1034	18.62988	21.91318
CCPI	30.99966	22.71138	13.02553	75.34519	17.46433	56.33716

Arsenic (As): Mean = 115.76 ppm; CV% = 149% –Elevated As suggests pervasive sulfide mineralization, possibly associated with arsenopyrite (Boyle, 1979).

Copper (Cu): High variability (CV > 90%) suggests spatially localized Cu-rich mineralization (Kesler, 1997).

Zinc (Zn) and Lead (Pb): CV% above 90% confirms the polymetallic nature of mineralization zones (Cooke et al., 2000).

Barium (Ba): With a mean of 2164 ppm and CV% = 88%, Ba appears as a marker for barite-related metasomatic zones (Heinrich, 2007).

Elements with lower CV%, such as Al and Be, appear more uniformly distributed, reflecting background lithological characteristics rather than hydrothermal inputs (Rollinson, 1993).

These statistics support the zonation interpreted in PCA and clustering analyses. Elements with high CV% correspond with zones of known mineral enrichment, reinforcing their use as exploration indicators. This quantitative summary validates the significance of elemental variability across alteration and mineralization zones, enhancing the reliability of the integrated classification model.

In summary, the results confirm a progression from magmatic-sulfide processes through metasomatic alteration to polymetallic mineralization. Geochemical indicators (Cu-Ag-Pb index, AI, CCPI) proved effective in zoning and classification, forming a solid foundation for predictive modeling in hydrothermal systems.

4. Discussion

This study applied an integrated geochemical, statistical, and machine learning framework to decipher complex mineralization and alteration processes within the Bektakari-Bnelikhevi ore knot, a representative segment of Georgia's hydrothermal system. By combining traditional analytical techniques with advanced data-driven tools, we provide a multi-scale interpretation of hydrothermal evolution, contributing to both scientific understanding and applied exploration strategies (Rollinson, 1993; Heinrich, 2007).

The strong elemental correlations (e.g., V-Sc, Mo-W, S-V) indicate coupled behavior of siderophile and chalcophile elements under high-temperature hydrothermal regimes, a hallmark of porphyry and skarn environments (Evans, 1993; Sillitoe, 1993). These associations reinforce the role of magmatic fluids as primary metal carriers, undergoing successive fractionation and interaction with host rocks (Kesler, 1997). The negative correlations, particularly between K and Na, mirror progressive metasomatic alteration –transitioning from sodic to potassic assemblages– as temperature, pH, and fluid composition evolve (Pirajno, 2009). This trend is further confirmed through AI and CCPI discrimination, which robustly segment the hydrothermal system into sericitic, intermediate, chloritic, and Na-Ca alteration zones, each reflecting distinct physicochemical conditions (Pearce et al., 1984).

PCA reveals that PC1 (49.9%) captures mineralization-related variance, driven by Zn, S, V, and Sc, suggesting early high-temperature sulfide deposition (Jolliffe and Cadima, 2016; Cooke et al., 2000). PC2 (15.8%) differentiates alkali- and barite-dominated metasomatic processes, as seen through high Na, K, and Ba loadings (Heinrich, 2007). Clustering analysis further isolates geochemical groups into paragenetic stages: early magmatic-sulfide (V, S, Zn), mid-stage feldspar alteration (Na, K), and late-stage polymetallic enrichment (Cu, Ag, Pb, Au). This sequence aligns with hydrothermal mineral system models and demonstrates that dimensionality reduction and clustering can effectively reconstruct mineralizing chronologies.

The zonal analysis of metal concentrations reinforces a spatial and temporal control on mineralization. High-Au zones feature elevated Zn, Cu, Pb, and Fe-As, characteristic of proximal, sulfide-rich mineralization. Medium-Au zones, with enriched Mo and W, likely represent distal or structurally isolated hydrothermal pulses, perhaps reflecting episodic fluid input or cooling events (Taylor and McLennan, 1985). The Low-Au zones correspond to geochemical

background, with low metal enrichment and minimal hydrothermal overprint. These insights provide direct evidence for phase-specific mineralization and fluid evolution (Boyle, 1979).

The dual-axis ratio analysis (log Cu/Mo vs. log K/Na) reveals meaningful patterns of vertical and lateral mineral control. Epithermal zones concentrate along High-K pathways, reflecting fluid ascent and lateral dispersion, while deep porphyry systems plot along Cu-rich vertical gradients, indicative of root-zone mineralization (Wilkinson, 2001). This differentiation underscores the need to integrate geochemical data with structural geology to map fluid pathways and alteration halos, especially in complex volcanic-arc settings (Richards, 2003).

Feature importance analysis emphasizes the effectiveness of composite predictors such as the Cu-Ag-Pb Index and Metallicity Factor. These indices encapsulate the cumulative behavior of ore-related elements and outperform individual predictors in classification tasks. The model's precision and recall were near-perfect in sericitic and Na-Ca zones, confirming the stability and geochemical isolation of these environments. Conversely, the model underperformed in transitional zones (Recall ~0.38), highlighting a persistent challenge in distinguishing mixed geochemical signals. This limitation, noted by reviewers, underscores the necessity of deeper feature engineering and spatial validation (Ying, 2019).

The statistical summary (Section 3.7) illustrates significant variability (CV% > 90%) in key ore metals, affirming their diagnostic relevance. Silver's CV of 192%, for example, reflects high spatial heterogeneity typical of epithermal systems. The broad distribution of As, Zn, Pb, and Cu supports their use as geochemical vectors in polymetallic mineral exploration (Boyle, 1979; Rollinson, 1993;). By correlating statistical dispersion with mineralogical zoning, we create a robust interpretive framework for resource delineation.

Model Limitations and Areas for Improvement. While the machine learning approach demonstrated high performance in clearly zoned environments, it struggled to accurately classify samples from transitional and background zones. Several factors likely contribute to this weakness:

Geochemical Overlap: Transitional zones inherently possess mixed characteristics of adjacent alteration types, leading to overlapping signatures that blur classification boundaries.

Limited Label Balance: The number of training samples for transitional and background zones was lower, creating a class imbalance that biases the model toward more dominant classes.

Insufficient Feature Diversity: Although indices like Cu-Ag-Pb and Metallicity Factor were effective, additional predictors, such as trace elements, REEs, and mineralogical indicators, could enhance resolution.

Lack of Spatial and Structural Integration: Geochemical data alone may not capture the full complexity of structural boundaries. Incorporating structural parameters or spatial modeling may provide clearer contextual cues for classification (Goovaerts, 1997).

To address these challenges, future iterations of the model should:

Implement sampling strategies to balance class representation.

Apply semi-supervised or ensemble learning methods that can manage noisy or ambiguous data.

Utilize domain knowledge-informed feature engineering, especially to capture subtle zonal transitions.

Scientifically, this study contributes to the growing body of research advocating for hybrid workflows in mineral exploration–those that combine geological reasoning with algorithmic discovery. Practically, the Cu-Ag-Pb Index and Metallicity Factor can be operationalized as exploration vectors in porphyry-epithermal terrains. AI and CCPI remain cost-effective proxies for alteration mapping, and their validation here supports their broader applicability. Structural-geochemical ratios (Cu/Mo, K/Na) offer additional

interpretive value and should be explored in other metallogenic provinces (Burnham, 1979; Sillitoe, 2010).

In summary, the integration of statistical, geochemical, and machine learning methodologies presented in this study provides a powerful toolkit for understanding mineralization controls. While the framework is highly effective in geochemically distinct zones, addressing its limitations in transitional environments is critical for broader application. With further refinement, this approach holds strong potential for guiding mineral exploration and deepening geochemical insights across varied hydrothermal systems.

5. Conclusion

This study developed and applied an integrated geochemical and machine learning framework to investigate hydrothermal alteration and mineralization processes in the Bektakari-Bnelikhevi ore knot. Grounded in a combination of multivariate statistical analysis, supervised classification, and geochemical index modeling, the research offers a replicable approach to understanding the structural, spatial, and compositional complexities of mineralized systems.

The correlation matrix revealed distinct elemental associations that are diagnostic of key hydrothermal processes: V-Sc and Mo-W point to high-temperature magmatic inputs; K-Na and K-Ca capture metasomatic overprinting. These patterns were validated by alteration indices (AI and CCPI), which effectively mapped the progression from early Na-Ca alteration to late sericitic and polymetallic enrichment zones.

Principal Component Analysis identified two orthogonal geochemical trends-magmatic-sulfide mineralization and alkali-driven metasomatism, each capturing a major axis of geochemical variability. Hierarchical clustering and unsupervised algorithms further refined zonation, reinforcing the evolutionary sequence of hydrothermal fluid activity. Distinct mineralization styles were observed across gold-bearing zones, with high-Au zones hosting elevated Zn, Cu, Pb, and Fe-As concentrations.

Structural geochemical ratios, such as log(Cu/Mo) and log(K/Na), contextualize mineralization in terms of fluid flow control and vertical-lateral partitioning of ore deposition. These relationships aligned with known tectono-magmatic settings and supported zoning interpretations based on alteration indices.

Machine learning classification achieved high precision and recall in sericitic and Na-Ca zones, driven by the predictive strength of the Cu-Ag-Pb Index and Metallicity Factor. However, as highlighted in the Discussion, model performance dropped significantly in transitional and background zones. These weaknesses are attributed to overlapping geochemical signals, class imbalance, and the exclusion of structural variables. This insight emphasizes the need for feature expansion, balanced sampling, and the incorporation of more geological context into future modeling workflows.

Statistical analysis of the dataset supported these findings, revealing high coefficients of variation for ore-related elements like Ag, Cu, Pb, and Zn. These metrics confirmed strong geochemical heterogeneity and further justified the use of multidimensional approaches.

In conclusion, this research demonstrates the value of integrating geochemical knowledge with statistical rigor and machine learning for mineral exploration. While the current model performs well in discrete geochemical environments, addressing its limitations in transitional zones is essential for broader applicability. Future studies should incorporate additional variables, more geological, geochemical data, and adaptive learning algorithms to enhance predictive capacity. The methodological structure proposed here serves as a strong foundation for mineral exploration in complex geological environments and provides a clear pathway for improving decision-making in resource targeting and management.

Acknowledgment

The authors express their deepest gratitude to Professor Vazha Geleishvili for his invaluable guidance and scholarly insights throughout this research. His expertise in geological sciences provided critical direction at various stages of the study.

We are also profoundly thankful to Dr. David Hyndman, Dr. Leili (Fatemeh) Izaditame, and Amin Tamadoni of the University of Texas at Dallas for their generous support and expert consultation in the fields of machine learning and data analysis, which substantially enriched the methodological framework of this work.

Furthermore, we extend our appreciation to mining company Rich Metals Group (RMG) for their provision of essential analytical materials, without which this study would not have been possible.

References

Adamia, Sh., Zakariadze, G., Sadradze, N., Beridze, T., Khutsishvili, S. 2017. Impact of Magmatism on the Geodynamic Evolution of Southern Georgia on the Example of the Lesser Caucasus Artvin-Bolnisi Block. EGU General Assembly 2017, Geophysical Research Abstracts, 19, EGU2017-6894.

Barton, M. D., Johnson, D. A. 1991. Evolution of magmatic to hydrothermal processes in porphyry copper systems. Journal of Geochemical Exploration 42(1), 73-99.

Boyle, R. W. 1979. The geochemistry of gold and its deposits. Geological Survey of Canada Bulletin 280, 584.

Breiman, L. 2001. Random forests. Machine Learning 45(1), 5-32.

Burnham, C. W. 1979. Magmas and hydrothermal fluids. Geochemistry of Hydrothermal Ore Deposits 71-136.

Cooke, D. R., Hollings, P., Walshe, J. 2000. Giant porphyry deposits: Characteristics, distribution, and tectonic controls. Economic Geology 95(4), 773-802.

Evans, A. M. 1993. Ore geology and industrial minerals: An introduction. Blackwell Science Ltd.

Franklin, J. M., Gibson, H. L., Galley, A. G., Jonasson, I. R. 2005. Volcanogenic massive sulfide deposits. Economic Geology 100th Anniversary, 523-560.

Goovaerts, P. 1997. Geostatistics for natural resources evaluation. Oxford University Press.

Groves, D. I., Goldfarb, R. J., Gebre-Mariam, M., Hagemann, S. G., Robert, F. 1998. Orogenic gold deposits: A proposed classification in the context of their crustal distribution and relationship to other gold deposit types. Ore Geology Reviews 13(1-5), 7-27.

Gustafson, L. B., Hunt, J. P. 1975. The porphyry copper deposit at El Salvador, Chile. Economic Geology 70(5), 857-912.

Heinrich, C. A. 2007. Fluid-fluid interactions in magmatic-hydrothermal ore formation. Reviews in Mineralogy and Geochemistry 65(1), 363-387.

Hedenquist, J. W., Lowenstern, J. B. 1994. The role of magmas in the formation of hydrothermal ore deposits. Nature 370(6490), 519-527.

Huston, D. L., Large, R. R., Cooke, D. R. 1995. The geochemical signature of volcanic-hosted massive sulfide mineralization and hydrothermal alteration. Australian Journal of Earth Sciences 42(6), 713-725.

Ishikawa, Y., Sawaguchi, T., Horiuchi, M., Harada, Y. 1976. Delineation of prospecting targets for Kuroko deposits based on modes of volcanism of underlying dacitic rocks. Mining Geology 26(5), 105-117.

Jolliffe, I. T., Cadima, J. 2016. Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374(2065), 20150202.

Kesler, S. E. 1997. Metallogenic evolution of the northern Andes. Economic Geology 92(5), 509-522.

Large, R. R., Gemmell, J. B., Paulick, H., Huston, D. L. 2001. The alteration box plot: A simple approach to understanding the relationship between alteration mineralogy and lithogeochemistry associated with volcanic-hosted massive sulfide deposits. Economic Geology 96(5), 957-971.

Lowell, J. D., Guilbert, J. M. 1970. Lateral and vertical alteration-mineralization zoning in porphyry ore deposits. Economic Geology 65(4), 373-408.

Mederer, J., Moritz, R., Zohrabyan, S., Ulianov, A., Melkonyan, R., Fikre, Y. 2014. Base and precious metal mineralization in Middle Jurassic rocks of the Lesser Caucasus: A review of geology and metallogeny and new data from the Kapan, Alaverdi and Mehmana districts. Ore Geology Reviews 58, 185-207.

Meinert, L. D. 1992. Skarn and porphyry mineralization in the Far Southeast deposit, Philippines. Economic Geology 87(5), 1143-1172.

Meinert, L. D., Dipple, G. M., Nicolescu, S. 2005. World skarn deposits. Economic Geology 100th Anniversary Volume, 299-336.

Nesbitt, H. W., Young, G. M. 1982. Early Proterozoic climates and plate motions inferred from major element chemistry of lutites. Nature 299(5885), 715-717.

Pearce, J. A., Harris, N. B. W., Tindle, A. G. 1984. Trace element discrimination diagrams for the tectonic interpretation of granitic rocks. Journal of Petrology 25(4), 956-983.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É. 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825-2830.

Pirajno, F. 2009. Hydrothermal processes and mineral systems. Springer Science and Business Media.

Powers, D. M. W. 2011. Evaluation: From precision, recall, and F-measure to ROC, informedness, markedness, and correlation. Journal of Machine Learning Technologies 2(1), 37-63.

Richards, J. P. 2003. Tectono-magmatic precursors for porphyry Cu-(Mo-Au) deposit formation. Economic Geology 98(8), 1515-1533.

Rollinson, H. R. 1993. Using geochemical data: evaluation, presentation, interpretation. Longman Group UK.

Saito, T., Rehmsmeier, M. 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE 10(3), e0118432.

Seedorff, E., Dilles, J. H., Proffett, J. M., Jr., Einaudi, M. T., Zurcher, L., Hedenquist, J. W., Stavast, W. J. A., Johnson, D. A., Barton, M. D. 2005. Porphyry deposits: Characteristics and origin of hypogene features. Economic Geology 100th Anniversary Volume, 251-298.

Sillitoe, R. H. 1993. Gold-rich porphyry copper deposits: geological model and exploration implications. Exploration and Mining Geology 2(4), 273-310.

Sillitoe, R. H. 2010. Porphyry copper systems. Economic Geology 105(1), 3-41.

Sosson, M., Rolland, Y., Müller, C., Adamia, S., Kangarli, T., Gasparyan, B., Galoyan, G., Sosson, A., Friedman, M., Vernant, P., Oberhänsli, R., Mosar, J. 2010. Subductions, obduction, and collision in the Lesser Caucasus (Armenia, Azerbaijan, Georgia): New insights. Geological Society, London, Special Publications 340(1), 329-352.

Sun, S.-S., McDonough, W. F. 1989. Chemical and isotopic systematics of oceanic basalts: Implications for mantle composition and processes. In A. D. Saunders and M. J. Norry (Eds.), Magmatism in the Ocean Basins, 313-345. Geological Society London, Special Publications.

Taylor, S. R., McLennan, S. M. 1985. The continental crust: its composition and evolution. Blackwell Scientific Publications.

Walker, E. W. 2015. Eurasian Geopolitics. UC Berkeley. Retrieved from https://eurasiangeopolitics.com

Wilkinson, J. J. 2001. Fluid inclusions in hydrothermal ore deposits. Lithos 55(1-4), 229-272.

Ying, Y. 2019. An overview of gradient boosting algorithms. Frontiers of Computer Science 13(1), 56-65.

Zhao, F. J., McGrath, S. P., Crosland, A. R. 1994. Comparison of three wet digestion methods for the determination of plant sulphur by inductively coupled plasma atomic emission spectroscopy (ICP-AES). Communications in Soil Science and Plant Analysis 25(3-4), 407-418.