Journal of Mining and Environment (JME) Vol. 18, No. 4, 2026, 1175-1193

Shahrood University of Technology

Journal of Mining and Environment (JME)

Journal homepage: www.jme.shahroodut.ac.ir

Iranian Society of Mining Engineering (IRSME)

An Innovative Approach to Mineral Resource Classification Based on Riemannian Clustering and Machine Learning in a Copper Deposit

Marco Antonio Cotrina-Teatino[1,4*], Jairo Jhonatan Marquina-Araujo[1,4], Jose Nestor Mamani-Quispe[2,4], Solio Marino Arango-Retamozo[1,4], Joe Alexis Gonzalez-Vasquez[3] and Kevin Daniel Rondo-Jalca[1,4]

1. Department of Mining Engineering, Faculty of Engineering, National University of Trujillo, Trujillo, Peru
2. Faculty of Chemical Engineering, National University of the Altiplano of Puno, Puno, Peru
3. Department of Industrial Engineering, Faculty of Engineering, National University of Trujillo, Trujillo, Peru
4. Laboratory for Mine Design and Planning – DIPLAMIN, National University of Trujillo, Trujillo, Peru

Article Info

Received 28 May 2025

Received in Revised form 27 June 2025

Accepted 6 September 2025

Published online 6 September 2025

DOI: 10.22044/jme.2025.16320.3174

Keywords

Mineral resource classification

Riemannian geometry

Geostatistics

Abstract

The classification of mineral resources significantly impacts mine planning, economic feasibility, and regulatory compliance. Despite its importance, such classification frequently depends on the subjective judgment of the Qualified Person (QP), owing to the absence of internationally standardized technical criteria for delineating resource categories. To mitigate this limitation, an innovative methodology integrating clustering based on Riemannian geometry with machine learning techniques was developed for mineral resource classification. A database of 5,654 composited samples from 185 diamond drill holes in a copper deposit in central Peru was utilized to classify 318,443 blocks. Copper grades were estimated through Ordinary Kriging (RMSE = 0.102; MAE = 0.069), generating geostatistical variables kriging variance, average distance to samples, and number of samples that served as input features for the classification. Clustering was performed using both classical KMeans and Riemannian KMeans, followed by spatial smoothing via XGBoost and Random Forest algorithms. Absolute coordinates were incorporated to address spatial discontinuities in classification outputs. The combination of the Riemannian model with Random Forest produced the highest classification performance, with a Silhouette index of 0.26 and a Davies-Bouldin index of 0.72. The resulting metal content was estimated at 4.24 Mt of copper at 0.44% grade (measured), 6.49 Mt at 0.34% Cu (indicated), and 7.68 Mt at 0.32% Cu (inferred), demonstrating close alignment with QP estimates while exhibiting improved spatial coherence. In summary, the Riemannian-based approach outperformed classical KMeans and conventional classification methods, providing a more robust, objective, and globally consistent alternative.

1. Introduction

Mineral resource classification is a fundamental stage in the evaluation of mining projects, as it provides the technical foundation for economic, engineering, and environmental decision making. Mineral resources are defined as naturally occurring concentrations of solid materials in the Earth's crust with potential economic value [1]. The international framework established by the Committee for Mineral Reserves International Reporting Standards (CRIRSCO) [2] classifies mineral resources into three categories Measured, Indicated, and Inferred based on the level of geological knowledge and the reliability of supporting data. Measured resources represent the highest level of confidence, whereas Inferred resources are based on limited sampling and indirect evidence. This classification system is adopted by several jurisdictions, including Australia [3], Canada [4], South Africa [5], Europe [6], the United States [7], Chile [8], and China under standard GB/T 17766-2020 [9, 10], among others.

Corresponding author: mcotrinat@unitru.edu.pe (M. Cotrina-Teatino)

Despite the existence of these regulatory frameworks, mineral resource classification remains a process subject to considerable expert judgment, as none of the reporting codes explicitly define the specific methods, parameters, or threshold values to be applied [11–13]. In practice, the judgment of the Qualified Person (QP) plays a decisive role, and decisions are based largely on their experience and contextual interpretation of the deposit's geometry. This reliance has led to significant variations in the applied criteria across projects, directly affecting the estimated tonnage, grade, and metal content by category [14–19]. Traditional approaches typically consider factors such as geological continuity, sampling density, grade variability (measured using variograms or covariance matrices), and the expected confidence level within the block model [20, 21]. However, when applied to complex multivariate systems, these methods can result in spatial discontinuities and outputs that are sensitive to manual segmentation and subjective boundaries [22–25].

In response to these challenges, a variety of alternative methodologies have been explored to address the limitations of conventional classification techniques. Early efforts focused on applying smoothing algorithms such as moving window filters to reduce discontinuities in block models [22]. More recently, machine learning-based models have gained traction due to their ability to model non-linear relationships and capture complex multivariate patterns in large databases. For instance, Support Vector Machines (SVM) with kernel functions have been applied to reclassify blocks based on spatial and statistical features [14], while artificial neural networks have been combined with clustering algorithms such as K-prototypes [26, 27]. Ensemble models like Random Forest have also shown promising results in the robust classification of geological blocks [28]. In parallel, advanced geostatistical techniques have been employed to better represent uncertainty and spatial continuity in categorical models. These include truncated Gaussian simulations, concentration–volume (C–V) fractal models, and rotational Truncated Bi-Gaussian simulations (TBSIM) [29, 30]. A pioneering fuzzy logic-based approach has also been applied in mineral resource classification [31]. While these methodologies have significantly improved the accuracy and objectivity of classification, they still face limitations in preserving the geometric structure of data in high-dimensional spaces with complex variable interdependencies.

In this context, clustering techniques based on numerical variables have become relevant for mineral resource classification. The K-Means algorithm, known for its versatility and computational efficiency, has been widely applied to key geostatistical variables such as kriging variance, average distance to samples, and number of samples per block. These variables capture fundamental aspects of geological knowledge and sampling quality. Recent studies have demonstrated that K-Means can effectively identify internal geological classification domains, particularly when applied to multivariate block models, as shown in the works of [32–34]. However, by operating under the assumption of a Euclidean space, this type of clustering may fail to adequately capture the underlying structural relationships between variables, thereby limiting its ability to preserve spatial coherence in complex block models.

In recent years, Riemannian geometry has emerged as a powerful mathematical tool for data analysis in non-Euclidean spaces, particularly in contexts where the internal structure of data plays a critical role. This approach recognizes that certain data types such as symmetric positive definite (SPD) covariance matrices cannot be accurately analyzed using traditional Euclidean metrics, as these do not preserve the intrinsic properties of the space in which the data reside. Instead, Riemannian geometry provides a framework based on differentiable manifolds and geodesic distances, allowing the preservation of the data space topology and enabling more coherent clustering, classification, and dimensionality reduction methods [35]. Successful applications have been documented in fields such as neuroimaging, for EEG signal classification [36], computer vision and electronics [37], as well as in mining, where it has been used for the segmentation of geological domains in multivariate environments [38].

Despite progress in this field, there remains a gap in the systematic application of Riemannian geometry-based methods for the direct classification of mineral resources. While its use in the spatial segmentation of geological domains has been explored, its implementation for assigning resource categories measured, indicated, or inferred based on geostatistical estimation variables has not yet been addressed [38]. This work seeks to fill that gap by proposing an innovative approach that combines block clustering via Riemannian K-Means, representing each block with local covariance matrices, followed by supervised smoothing using algorithms such as XGBoost and Random Forest to enhance spatial coherence in mineral resource classification.

This article is structured as follows: first, the database and methodology are described in detail, including both the Riemannian model and the supervised smoothing process. Then, the method is applied to a copper deposit located in Peru, with results presented and compared against conventional techniques. Finally, the paper concludes with key findings and proposed directions for future research.

2. Methodology

2.1. Study area: Geological and mineralization setting

The studied deposit is an open-pit copper–molybdenum system located in the central Andes of Peru, at approximately 4,600 meters elevation. The lithological sequence is geologically complex and comprises five main units: magnetite-bearing skarn, granodioritic intrusive rocks, dacitic porphyry bodies, calcareous sedimentary formations, and a volcanic sequence correlated with the Catalina Formation. Copper and molybdenum mineralization is distributed throughout lithologies, but preferentially concentrated in the magnetite skarn and granodiorite, which act as the principal hosts for high-grade mineralization. In contrast, the dacitic porphyry, calcareous units, and volcanic rocks exhibit lower and more variable mineral concentrations.

The mineralizing system reflects a transitional porphyry–skarn environment, with mineralization occurring in multiple forms, including disseminated sulfides, veinlet stockworks, and massive sulfide replacement zones. These styles are primarily associated with hydrothermal processes and are spatially controlled by fracture systems, lithological contacts, and structural discontinuities that facilitated focused fluid flow and metal deposition.

2.2. Database description

This study is based on a geological database comprising 5,654 composited samples obtained from 185 diamond drill holes completed in a copper deposit located in central Peru. The drill holes are spaced at an average interval of 30 meters and extend to an average depth of approximately 480 meters. Compositing was performed following lithological boundaries, using fixed intervals of 15 meters to ensure consistency and homogeneity across the database. Figure 1 presents the spatial distribution of the drill holes along with corresponding copper grade values. Figure 2 displays the histogram of copper grade distribution, highlighting a mean grade of 0.43% Cu within the composited database.

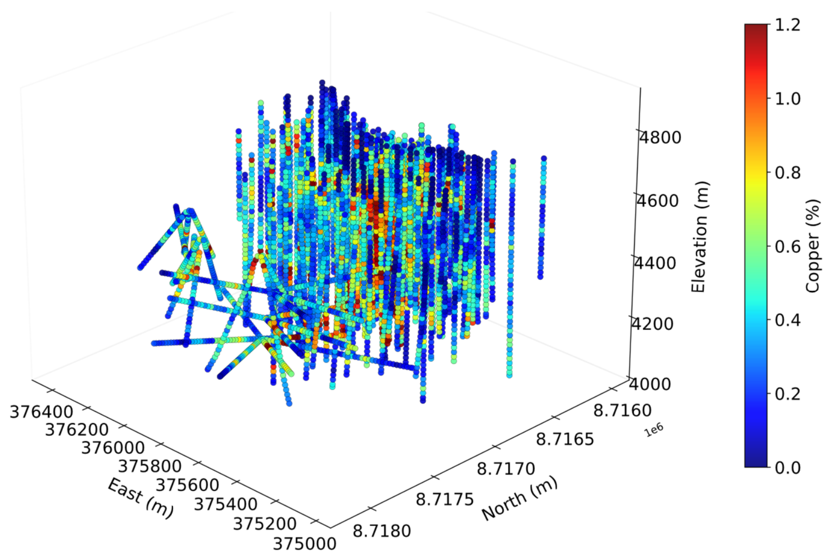


Figure 1. Spatial distribution of diamond drill holes and copper grade values

2.3. Variables used in mineral resource classification

The three-dimensional block model developed for this research covers a volume of 1,640 m (E–W), 2,200 m (N–S), and 920 m vertically. Regular blocks of 20 meters per side were used, resulting in a grid of 82 × 110 × 46 blocks, for a total of 414,920 spatial blocks. Copper grades were estimated using the Ordinary Kriging (OK) method, implemented via the Stanford Geostatistical Modeling Software (SGeMS) [39, 40]. Directional variograms were constructed for copper grade along the principal geological axes (0°, 45°, and 90°), with ranges varying from 150 to 200 m. The selected covariance function was spherical, with maximum, intermediate, and minimum ranges of 250 m, 130 m, and 115 m, respectively. The estimation was constrained within a search ellipsoid defined by radii of 450 m (maximum), 250 m (intermediate), and 200 m (minimum). Under these parameters, estimated values were obtained for 318,443 blocks, representing the mineralized portion of the deposit. To validate the consistency of the model, estimated values were compared against actual grades recorded in drill hole samples using a KD-tree algorithm, which efficiently matched each real sample to its nearest block [41]. The resulting evaluation metrics were: RMSE = 0.102, MAE = 0.069, and R^2= 0.875, indicating a robust fit of the estimated model (see Figure 3).

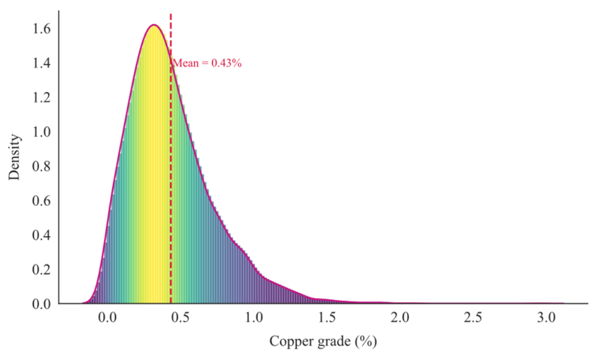


Figure 2. Probability density distribution of copper grades in the composite sample database

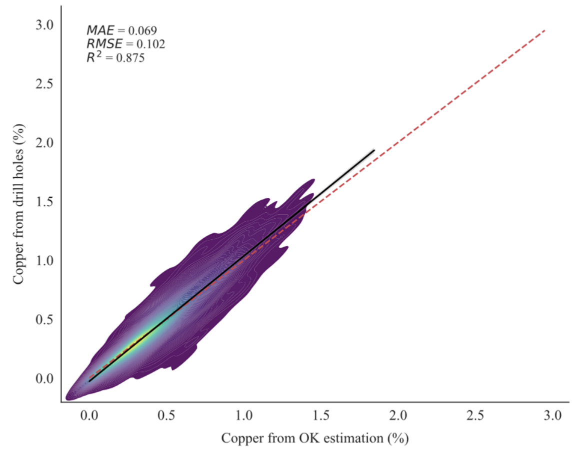


Figure 3. Comparison between estimated and actual copper grades using KD-tree validation

Based on this model, the variables used for mineral resource classification were derived: Kriging Variance (VOK), Average Distance to Samples (SD), and Number of Samples (NS). Table 1 presents the descriptive statistics of these variables for the complete set of estimated blocks. The kriging variance generally exhibited low values (mean value of 0.07), while the average distance to samples reached 231.91 m. The number of samples per block ranged from 10 to 200, with an average of 124. The spatial distribution of these variables is shown in Figure 4.

Table 1. Descriptive statistics of geostatistical variables used in the classification

Variable	Mean	Std	Min	Max	Variance	Skewness	Kurtosis
VOK	0.07	0.03	0.00	0.15	0.00	-0.54	-0.63
SD	231.91	59.19	114.27	433.19	3,503.18	0.51	-0.11
NS	124	70.99	10.00	200.00	5,040.28	-0.22	-1.57

2.4. Mineral resource classification using Euclidian and Riemannian clustering

Mineral resource classification was conducted using two clustering approaches: (i) the classical K-Means algorithm based on Euclidean distance, and (ii) a Riemannian geometry-based method operating in the space of symmetric positive definite (SPD) matrices. Figure 5 presents the methodological workflow adopted for the Riemannian-based classification. The models were developed in Python 3.11.11 using Jupyter Notebook as the working environment, and libraries such as scikit-learn, pyriemann, and pandas.

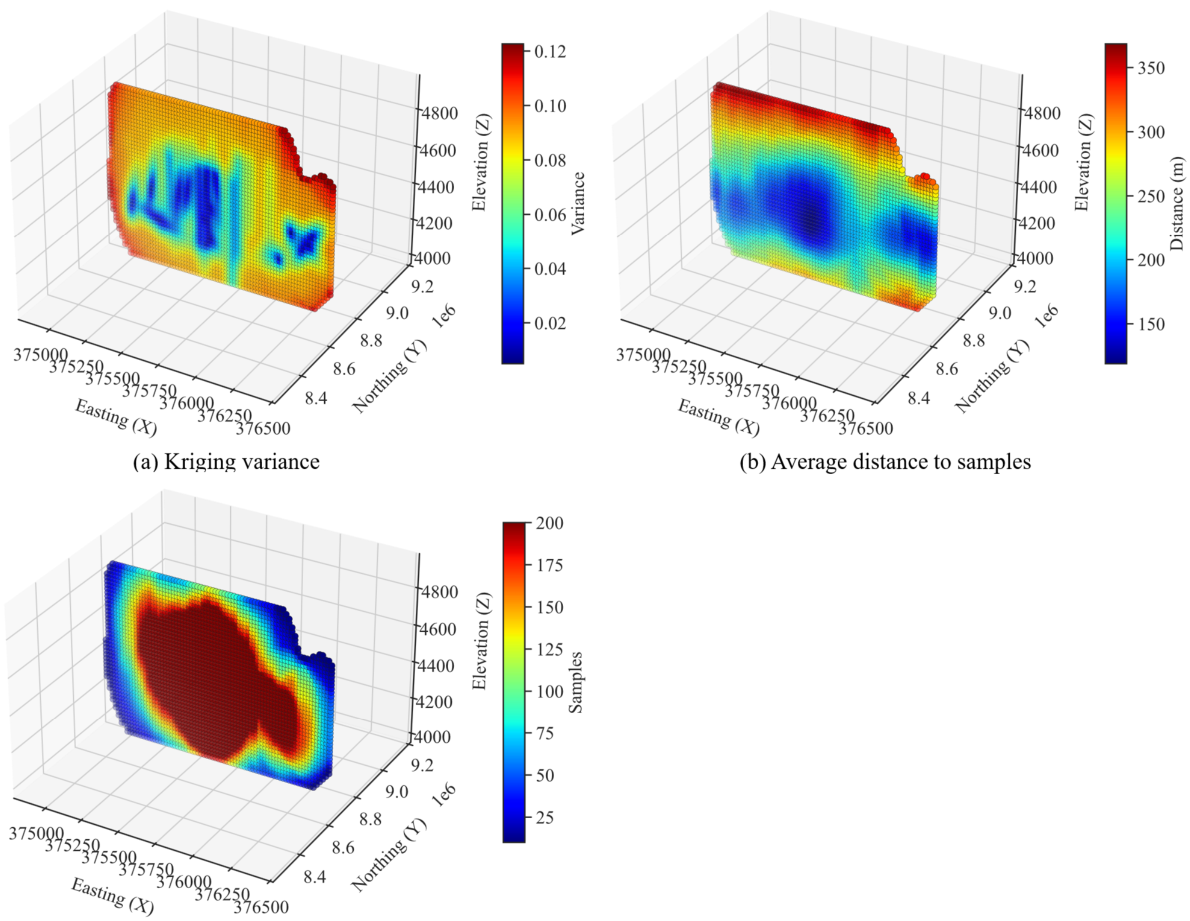
The QP's classification criteria were based exclusively on quantitative variables. In the univariate case, given a continuous variable X, two thresholds T_1 and T_2 were defined such that $T_1 < T_2$, and the classification was established as follows (Equation (1)).

$$\text{Measured: } X \leq T_1;\ \text{Indicated: } T_1 > X \leq T_2 \text{ and Inferred: } X > T_2 \tag{1}$$

For the multivariate case $X = (X_1, X_2, \dots, X_N)$, thresholds (T_{i1}, T_{i2}) were defined for each dimension as shown in Equation (2).

$$\text{Measured: } X_i \le T_{i1}\ \forall\, i;\ \text{Indicated: } T_{i1} < X_i \le T_{i2}\ \forall\, i \text{ and Inferred: } X_i > T_{i2}\ \forall\, i \tag{2}$$

In the multivariate case, the classification rule is applied jointly across all variables. A block is classified as "Measured" only if all variables satisfy the most restrictive thresholds ($X_i \le T_{i1}\ \forall\, i$); "indicated" if all variables fall within intermediate ranges ($T_{i1} < X_i \le T_{i2}\ \forall\, i$); and "inferred" if none of the variables meet the criteria for the previous categories. This ensures that each block is assigned to a single and unambiguous category, consistent with the QP's classification logic.



(a) Kriging variance

(b) Average distance to samples

(c) Number of samples

Figure 4. Representation of the geostatistical variables used for mineral resource classification

2.5. Classification using classical (Euclidean) K-Means

The classical K-Means method was applied as an unsupervised clustering technique in the Euclidean space $\mathbb{R}^n$, aiming to identify subsets of blocks with similar properties. The procedure involved the following steps:

Step 1. Data preparation: The selected numerical variables were transformed using the PowerTransformer in order to stabilize variance, correct skewness, and standardize scales, thereby preventing distortions in distance calculations [32, 42, 43].

Step 2. Model optimization: The objective function minimizes the sum of squared distances to the centroid of each cluster, as expressed in Equation (3).

$$J = \sum_{l=1}^{k} \sum_{x_i \in C_l} \|x_i - \mu_l\|^2 \quad (3)$$

where, C_l is the set of blocks assigned to cluster l, and μ_l is the Euclidean centroid.

Step 3. Category assignment: The three resulting clusters were assigned to the Measured, Indicated, and Inferred categories based on the statistical ordering of the centroids and their consistency with geological criteria.

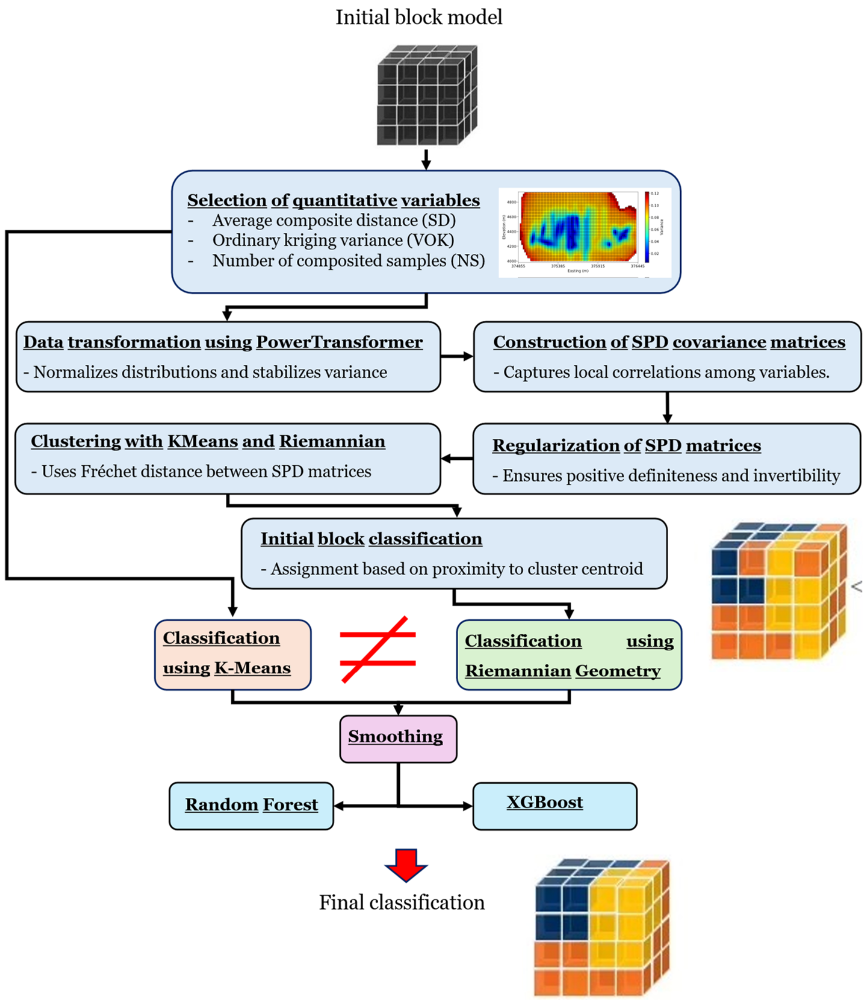


Figure 5. Methodological workflow for the Riemannian geometry-based classification model, including input data, clustering, and supervised smoothing

2.6. Classification using Riemannian geometry

Unlike the previous approach, this method accounts for the inherent geometry of the space of symmetric positive definite matrices, employing Riemannian distances instead of Euclidean ones. Each block was represented by a local covariance matrix $C_i \in \mathbb{R}^{pxp}$, constructed from the statistical distribution of the variables (SD, VOK, NS) whitin a local neighborhood of spatially adjacent blocks. Specifically, a 3x3 covariance matrix was computed for each block using the values of its k-nearest neighboring, capturing localized multivariate variability and correlations within a spatial context.

Step 1. Construction of SPD matrices [44, 45]: For each block, the following was computed (Equation (4)).

$$C_i = Cov(X_i) \quad (4)$$

where, C_i is the covariance matrix X_i, and X_i is the feature vector of block i. The variable SD was included as a continuous feature, allowing its influence on local dispersion to be modeled.

Step 2. Regularization [38, 46, 47]: To ensure positive definiteness and matrix invertibility, a regularization procedure was applied as shown in Equation (5).

$$C_i^{SPD} = C_i + \in I \quad (5)$$

where, I is the identity matrix and $\in \ll 1$ is a small scalar (typically 0.001).

Step 3. Riemannian distance metric [38, 48]: The Riemannian metric between SPD matrices was employed, defined in Equation (6).

$$d_{Riemann}(C_1, C_2) = \left\| \log(C_1^{-1/2} C_2 C_1^{-1/2}) \right\|_F \quad (6)$$

where, log denotes the matrix logarithm and $\|.\|_F$ is the Frobenius norm, defined as the square root of the sum of the squares of all matrix elements. This norm measures the “magnitude” of a matrix analogously to the Euclidean norm for vectors.

Step 4. Cluster centroid via Frechet mean: In this non-linear space, the cluster centroid does not correspond to an arithmetic mean but to the Fréchet mean, which minimizes the sum of squared Riemannian distances (Equation (7)).

$$\bar{C}_l = \underset{C}{\arg min} \sum_{i=1}^{n_l} d_{Riemann}(C_i, C)^2 \quad (7)$$

where, $\bar{C}_l$ represents the Frechet mean of cluster l, and n_l is the number of blocks in that cluster. The Frechet mean is a generalization of the mean to non-Euclidean spaces, respecting the intrinsic geometry of the SPD manifold.

Step 5. Final block classification: Each block is assigned to the category corresponding to the cluster whose centroid matrix $\bar{C}$ is the closest in terms of Riemannian distance. The final assignment yields a classification into the conventional categories of Measured, Indicated, and Inferred.

With respect to the classification methods, Table 2 presents the hyperparameters used for the mineral resource classification models.

Table 2. Hyperparameters used in the mineral resource classification models

Model	Purpose	Hyperparameter	Value
Classical K-Means	Unsupervised clustering in Euclidean space	n_clusters	3
		init	“k-means++”
		n_init	10
		random_state	17,276,365
		max_iter	300
		tol	1e-4
Riemannian K-Means	Unsupervised clustering in the SPD matrix space	n_clusters	3
		metric	“riemann"
		init	“random”
		max_iter	100
		tol	1x10-9
		random_state	42
		n_init	1

2.7. Classification smoothing using machine learning models

To increase the spatial coherence of the block model and enhance the quality of the classification produced by the clustering algorithms (both classical and Riemannian K-Means), a supervised smoothing stage was implemented using machine learning models. This phase focused on using the absolute spatial coordinates of each block east (X), north (Y), and elevation (Z) as independent variables, and the previously assigned category (Measured, Indicated, Inferred) as the dependent variable.

The fundamental assumption is that the category assigned to a mining block (y_i) should be spatially correlated with its immediate neighbors. Instead of using geostatistical features or grade values as inputs, the absolute spatial coordinates of each block, as defined in Equation (8).

$$X_i = [x_i, y_i, z_i] \ and \ y_i \in \{Measured, Indicated, Inferred\} \quad (8)$$

This allows the model to learn implicit spatial patterns in the category distribution, promoting continuous classification across space and correcting erratic assignments resulting from the lack of inherent smoothness in clustering methods.

The XGBoost model was selected due to its ability to capture nonlinear relationships, computational efficiency, and robustness against overfitting [49–53]. Ensemble approaches such as Machine Learning have been shown to significantly improve classification performance in geological and geochemical applications, as demonstrated by Farhadi et al. [54, 55]. The Python implementation of XGBoost was used, configuring the model with three classes ($num_{class} = 3$) and the loss function as ($multi\text{:}softmax$) for direct multiclass classification output. The database was split into 80% for training and 20% for validation, ensuring reproducibility with $random_{state} = 42$. Spatial coordinates were normalized using $StandardScaler$ to stabilize variance and improve training efficiency. The loss function used was the multiclass logistic loss, formally defined in Equation (9).

$$L = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{C} y_{ij} lo\, g(\hat{y}_{ij}), \tag{9}$$

where n is the number of samples, C=3 is the number of classes, $y_{ij} \in \{0,1\}$ is the true label for sample i and class j, and $\hat{y}_{ij}$ is the predicted probability. This metric heavily penalizes incorrect high-confidence predictions, promoting better probabilistic calibration.

Additionally, the Random Forest algorithm was applied as an ensemble model based on decision trees. This approach was chosen for its effectiveness in correcting local overfitting and its capacity to capture spatial interactions without the need for complex transformations [56–60]. The spatial coordinates were used as inputs, normalized, and the database was split 80/20 for training and validation. The model was configured with 60 trees, a maximum depth of 9, and controlled split functions to reduce the complexity of individual trees. Table 3 summarizes the hyperparameters used in both supervised smoothing models. These configurations were selected based on internal training iterations.

Although spatial clustering methods such as K-Means with spatial constraints or spatial Gaussian mixture models allow for the direct inclusion of coordinates in the clustering process, we chose not to apply them in this study. Our objective was to classify the blocks based solely on the reliability of the geostatistical estimation (SD, VOK, NS), without imposing spatial continuity from the outset. Including coordinates at this stage may lead to overly smoothed results, potentially masking important geological boundaries. Instead, we applied spatial smoothing afterward using supervised models (XGBoost and RF), which allowed us to better control spatial continuity without losing the geological relevance of the classification.

2.8. Classification performance evaluation

To assess the quality of the classification generated by the models, both traditional classification metrics and specific indicators of spatial clustering quality were applied. First, the corresponding confusion matrix was constructed [14], comparing the categories generated by the classical and Riemannian K-Means models, as well as their smoothed versions, with the classification performed by the QP. Additionally, three key metrics were calculated to evaluate the clustering structure:

Silhouette score [14], which measures the internal consistency of each cluster in relation to its separation from other clusters. This indicator ranges from -1 to 1, where higher values reflect better clustering structure.

Calinski-Harabasz index, which quantifies the dispersion between and within clusters. A higher value suggests that the clusters are well separated and compact [61].

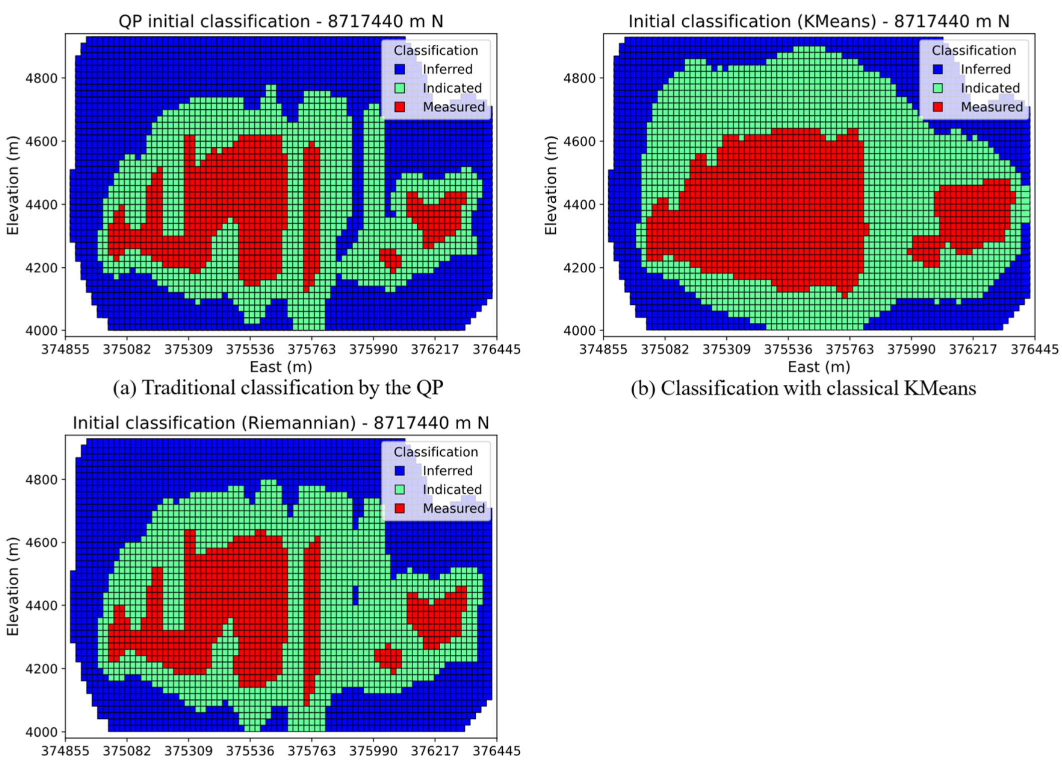
Davies-Bouldin index, which estimates the similarity between clusters by considering their dispersion and the distance between centroids. A lower value indicates better performance [62].

3. Results

The initial mineral resource classification reveals notable differences depending on the approach applied. The classification comprises three classes: Measured (red), Indicated (green), and Inferred (blue). The traditional interpretation performed by the Qualified Professional (QP) shows a more irregular distribution of the categories. In contrast, the KMeans clustering tends to generate more uniform zones. On the other hand, the Riemannian approach better captures transition patterns between categories, maintaining spatial coherence while respecting the geological shape of the mineralized body (see Figure 6). This aligns with findings by Riquelme et al. [38], who demonstrated that using differential geometry improves the structural representation of data in geologically complex deposits.

Table 3. Configuration parameters of the supervised machine learning models used to smooth the classification output based on spatial coordinates

Model	Purpose	Hyperparameter	Value
XGBoost	Multiclass on coordinates	objective	"multi:softmax"
		num_class	3
		random_state	42
		use_label_encoder	False
		eval_metric	"mlogloss"
		max_depth	5
		learning_rate	0.1
		n_estimators	100
		min_child_weight	5
		subsample	0.7
		colsample_bytree	0.7
		gamma	1.0
Random forest	Multiclass on coordinates	n_estimators	60
		max_depth	9
		random_state	42
		min_samples_split	20
		min_samples_leaf	10
		max_features	"log2"
		bootstrap	True



(a) Traditional classification by the QP

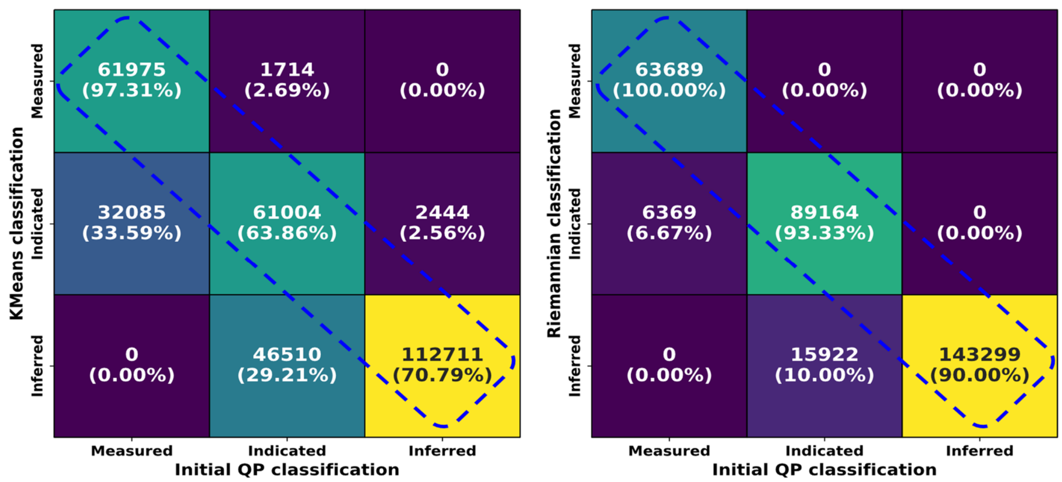
(b) Classification with classical KMeans

(c) Classification with Riemannian

Figure 6. Block classification comparison at section 8717440 m N

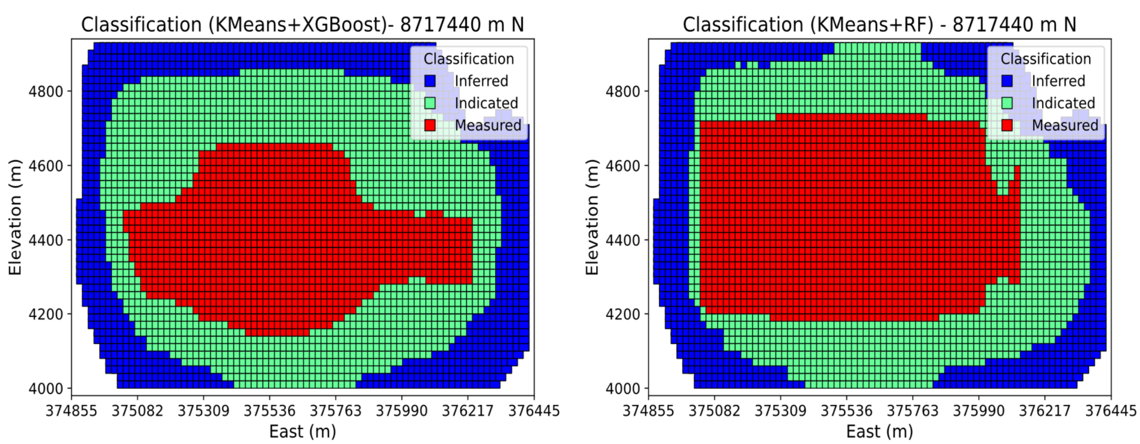
The degree of agreement between the automated classifications and the QP's reference classification was assessed using confusion matrices. In the case of K-Means, the "Measured" category achieved a high accuracy of 97.31%; however, substantial misclassification occurred in the "Indicated" category, with 29.21% of these blocks reassigned to the "Inferred" class. In contrast, the Riemannian approach demonstrated notable improvement, yielding 100% accuracy for the "Measured" category and 90% for the "Inferred" category, suggesting enhanced spatial consistency and alignment with the expert-based interpretation (see Figure 7). These findings highlight that conventional performance metrics may not fully capture the complexity of class boundaries, particularly in geologically transitional zones. The Riemannian metric aligns with the interpretative logic of the QP, as noted by Stephenson and Silva et al. [12, 21].

The use of supervised learning models enabled the post-processing and smoothing of the K-Means clustering results by incorporating spatial coordinates. The K-Means + XGBoost model produced a smoothed classification that retained the core structure of the mineralized body while enhancing lateral continuity in the "Indicated" zones. In comparison, the K-Means + Random Forest model resulted in a more regularized geometry, characterized by clearly delineated boundaries but with slight overgeneralization of the "Measured" category, which tended to extend into adjacent areas (see Figure 8). Given its sensitivity to local spatial patterns, Random Forest may overfit by disproportionately expanding the most confident classes. This type of overgeneralization in geospatial applications tends to optimize the objective function without respecting the structural heterogeneity of the deposit, as indicated by Cotrina et al. [28].



(a) QP classification vs KMeans classification (b) QP classification vs Riemannian classification

Figure 7. Confusion matrices comparing initial automatic classification (KMeans and Riemannian) with QP reference

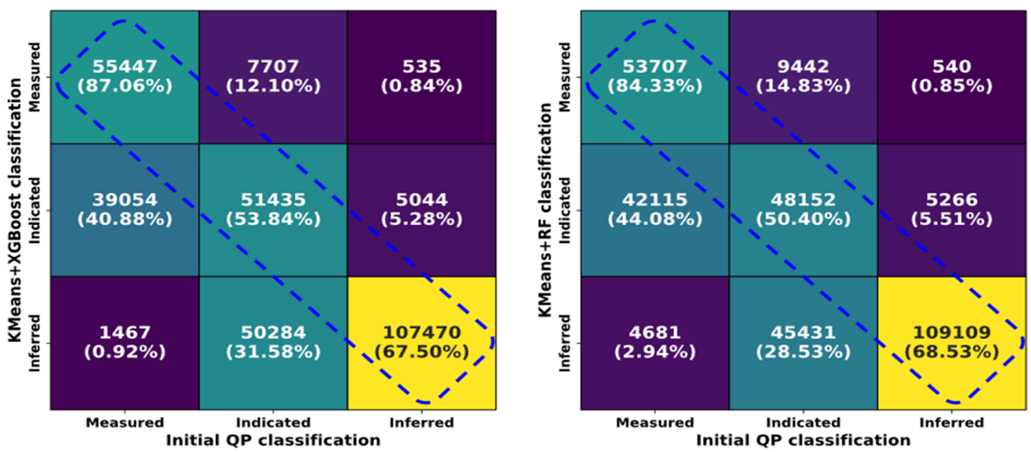


(a) KMeans+XGBoost (b) KMeans+Random forest

Figure 8. Classification smoothing of KMeans

Confusion matrices for the smoothed models reveal a trade-off between overall accuracy and spatial smoothness. The KMeans+XGBoost combination achieves better significance in the "Measured" class with an accuracy of 87.06%, while the "Indicated" and "Inferred" categories show some overlap. In contrast, the KMeans+RF model shows a slight decrease in global accuracy but a more homogeneous distribution, with an accuracy of 84.33% for “Measured” and a more conservative approach in the classification of “Inferred” (see Figure 9). This behavior has been analyzed by Cevik et al. [14], who report that ensemble models allow for more flexible separation but require careful calibration to avoid excessively conservative classification.

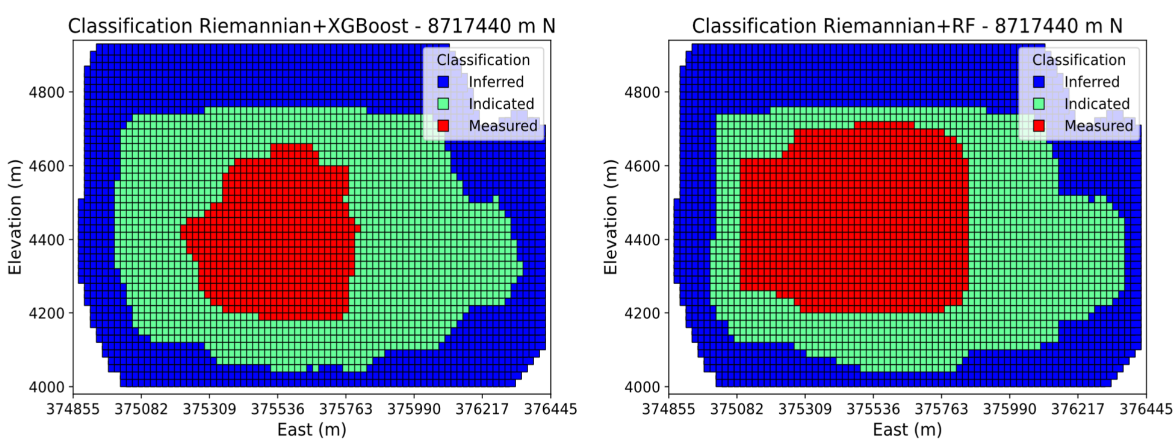
The Riemannian+XGBoost model produces a centralized distribution of the “Measured” body, with a well-defined concentric transition towards the “Indicated” and “Inferred” categories. This configuration suggests a coherent response from the model to spatial radial patterns of mineralization. In contrast, the Riemannian+RF model generates a broader expansion of the "Measured" domain, occupying a substantial portion of the mineralized core, which may reflect greater sensitivity to the spatial proximity of similar blocks (see Figure 10). This aligns with the ideas of Pennec et al. [35] on the importance of preserving the structure of the space when performing classification tasks in geoscientific contexts using Riemannian geometry.



(a) KMeans + XGBoost classification vs QP classification

(b) KMeans + RF classification vs QP classification

Figure 9. Confusion matrices for smoothed KMeans classification versus QP classification

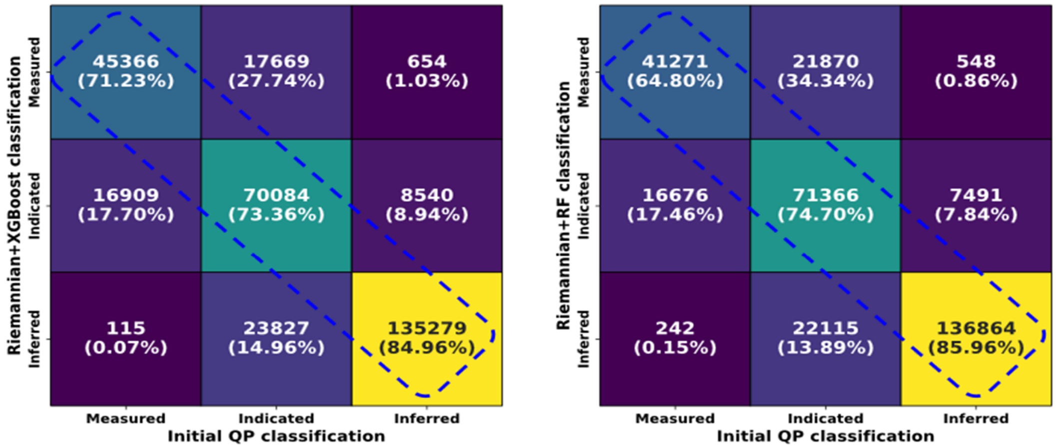


(a) Riemannian + XGBoost

(a) Riemannian + RF

Figure 10. Classification smoothing of Riemannian

The accuracy achieved by the supervised models was evaluated through confusion matrices cross-referenced with the QP classification as the reference (see Figure 11). The Riemannian+XGBoost model achieves a 71.23% agreement in the “Measured” category, 73.36% in “Indicated”, and 84.96% in “Inferred”, reflecting a balanced recovery and specificity. In comparison, the Riemannian+RF model reaches 64.80%, 74.70%, and 85.96%, respectively, indicating a slight loss in accuracy in the most reliable class in favor of better separation of the “Inferred” class.



(a) Riemannian + XGBoost classification vs QP classification

(b) Riemannian + RF classification vs QP classification

Figure 11. Confusion matrices comparing smoothed Riemannian classification to the QP interpretation

From a metric perspective, the comparison between the initial and smoothed models shows a significant impact on clustering quality. The Silhouette index increased from 0.14 (unsmoothed Riemannian model) to 0.26 with Riemannian+RF smoothing, which was the highest value of all models tested. Similarly, the Davies-Bouldin index was drastically reduced from 8.34 to just 0.72 with this last model, indicating more compact and well-defined clusters. In contrast, the classic KMeans model smoothed with XGBoost barely reached a Silhouette of 0.04 and a Davies-Bouldin of 17.74, reflecting a much less efficient structure. Table 4 summarizes these metrics, highlighting the superior geometric performance of the smoothed Riemannian approach with Random Forest.

Table 4. Summary of clustering evaluation metrics before and after smoothing

Model	Silhouette score	Calinski-Harabasz score	Davies-Bouldin score
Initial QP classification	0.15	17.25	8.01
Initial KMeans classification	0.05	7.14	26.96
KMeans + XGBoost classification	0.04	16.95	17.74
KMeans + RF classification	0.02	25.88	14.81
Initial Riemannian classification	0.14	22.37	8.34
Riemannian + XGBoost classification	0.25	16.20	2.26
Riemannian + RF classification	0.26	22.25	0.72

The QP estimation reports 1,171.88 Mt of mineral classified as "Measured" with 0.43% Cu, equivalent to 4.54 Mt of contained copper. The Riemannian model reproduces this category with 1,289.07 Mt (0.42% Cu), while KMeans overestimates it at 1,730.70 Mt, which increases the contained metal to 6.54 Mt, reflecting a more optimistic reclassification. For the "Indicated" class, all three models present comparable values, while in the "Inferred" category, the Riemannian model achieves a better approximation to the QP classification both in tonnage (2,636.70 Mt) and metal content (7.59 Mt Cu). These results are quantified in Table 5, demonstrating that the Riemannian geometry-based model fits better with the interpretative structure of the deposit. This overrepresentation in KMeans can induce optimistic biases in mining planning. As discussed by Madani and Owusu et al. [16, 30], overestimation of high-category blocks without solid statistical backing can compromise the economic viability of a project and its regulatory compliance, particularly in jurisdictions governed by CRIRSCO aligned codes, including NI 43-101 (Canada).

The Riemannian+RF model reduces the volume of blocks classified as "Measured" to 1,070.68 Mt, but raises the average grade to 0.44% Cu, which allows it to reach 4.24 Mt of contained copper, a value similar to the QP but with a more conservative base. In "Indicated", it reaches 2,122.46 Mt with 0.34% Cu (6.49 Mt), and in "Inferred" it exceeds 2,666.22 Mt, accumulating 7.68 Mt of copper, the highest value reported. This distribution suggests higher precision in categorizing low-confidence blocks. On the other hand, the KMeans+XGBoost model significantly increases the volume classified as "Measured" (1,765.81 Mt), but with a lower grade (0.42% Cu), indicating a tendency toward overclassification. This suggests that the Riemannian smoothed model with RF offers the best balance between geological accuracy and economic performance (see Table 6). This result suggests that the smoothed Riemannian model not only improves spatial representation but also optimizes resource estimation in lower-confidence categories.

Table 5. Comparison of initial resource classification results from QP, KMeans, and Riemannian models

Resources	Model	Number of blocks	Tonnage (Mt)	Copper (%)	Metal (Mt)
Measured	QP	63,689	1,171.88	0.43	4.54
	K-Means	94,060	1,730.70	0.42	6.54
	Riemannian	70,058	1,289.07	0.42	4.87
Indicated	QP	95,533	1,757.81	0.34	5.38
	K-Means	109,228	2,009.80	0.33	5.97
	Riemannian	105,086	1,933.58	0.34	5.92
Inferred	QP	159,221	2,929.67	0.32	8.44
	K-Means	115,155	2,118.85	0.31	5.91
	Riemannian	143,299	2,636.70	0.32	7.59

Table 6. Results of smoothed classification using KMeans and Riemannian models: tonnage, grade, and contained metal by category

Resources	Model	Number of blocks	Tonnage (Mt)	Copper (%)	Metal (Mt)
Measured	KMeans + XGBoost	95,968	1,765.81	0.42	6.67
	KMeans + RF	100,503	1,849.26	0.42	6.99
	Riemannian + XGBoost	62,390	1,147.98	0.45	4.65
	Riemannian + RF	58,189	1,070.68	0.44	4.24
Indicated	KMeans + XGBoost	109,426	2,013.44	0.33	5.98
	KMeans + RF	103,025	1,895.66	0.32	5.46
	Riemannian + XGBoost	111,580	2,053.07	0.33	6.10
	Riemannian + RF	115,351	2,122.46	0.34	6.49
Inferred	KMeans + XGBoost	113,049	2,080.10	0.31	5.80
	KMeans + RF	114,915	2,114.44	0.31	5.90
	Riemannian + XGBoost	144,473	2,106.30	0.32	6.07
	Riemannian + RF	144,903	2,666.22	0.32	7.68

The comparison of the grade-tonnage curves by mineral resource class provides additional evidence regarding the economic behavior of each model. In the "Measured" class, illustrated in Figure 12, it is observed that the KMeans model significantly overestimates the tonnage relative to the QP at low cut-off grades, while the Riemannian+RF approach shows a more conservative curve but with higher average grades, converging to similar QP values as the cut-off increases. In the case of the "Indicated" resources (Figure 13), the unsmoothed Riemannian model maintains an intermediate trajectory between KMeans' overrepresentation and the QP's restriction, although the smoothed versions tend to stabilize the tonnage curve and smooth the grade slope. On the other hand, the "Inferred" category, presented in Figure 14, shows the greatest variability range between models: the QP estimates the highest tonnage at low grades, but the smoothed curves, especially Riemannian+RF, capture greater selectivity at medium and high grades, with steeper grade curves.

The results show that the Riemannian approach offers clear advantages over traditional methods and supervised models based on Euclidean distances. By working with covariance matrices in a non-linear space, this approach allows for better classification of blocks while respecting the geological shape of the deposit and the spatial continuity between categories. Furthermore, when combined with algorithms like RF, it achieves a good balance between accuracy, geometric smoothness, and statistical stability. These advantages make the Riemannian model a solid alternative for reducing subjectivity in the classification of mineral resources [35, 38].

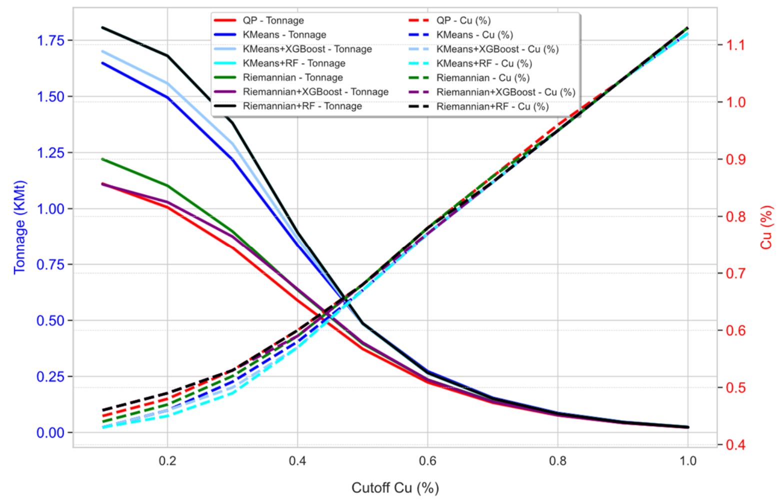


Figure 12. Tonnage vs. copper grade curves for the measured resource class across models

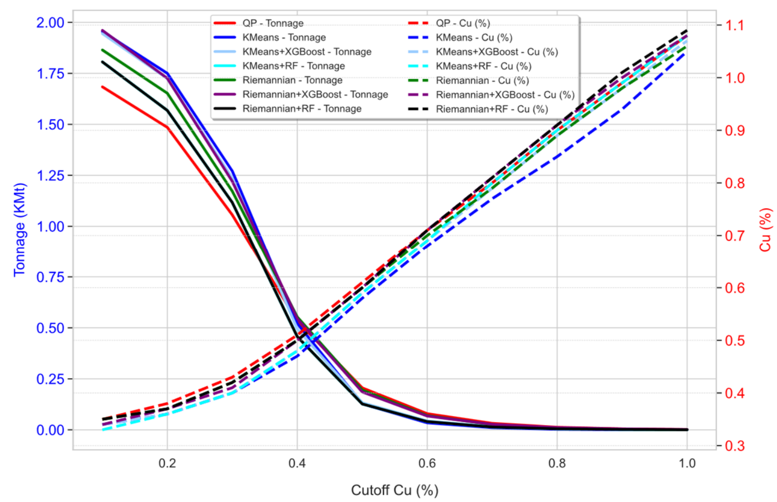


Figure 13. Tonnage vs. copper grade curves for the indicated resource class across models

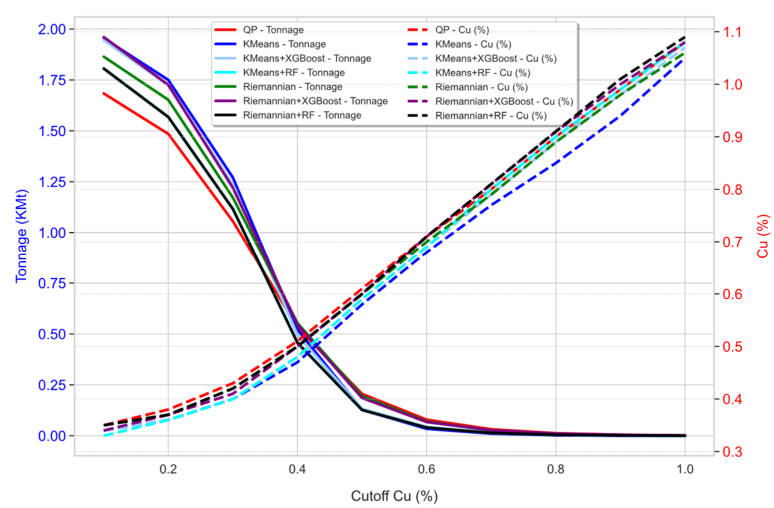


Figure 14. Tonnage vs. copper grade curves for the inferred resource class across models

4. Conclusions

This study proposed an innovative approach for the classification of mineral resources through the combined use of Riemannian geometry and machine learning models. The results obtained demonstrate that the Riemannian clustering model outperforms traditional methods, such as classic KMeans and the classification performed by the QP, both in terms of spatial coherence and categorical precision. By representing each block through local covariance matrices and applying distance metrics in non-Euclidean spaces, a more consistent segmentation in relation to the geological geometry of the deposit was achieved. Furthermore, the subsequent smoothing stage with RF allowed for refinement of spatial continuity without sacrificing the statistical quality of the model, achieving the highest values in indicators like Silhouette and the lowest Davies-Bouldin indices, with economic results that closely match the categories estimated by the QP. In summary, the proposed approach offers a solid and replicable alternative for improving objectivity and traceability in mineral resource classification. It is important to clarify that the proposed methodology is not intended to fully automate resource classification. Expert judgment remains essential for defining parameters and interpreting geological context. Rather than replacing this role, the method provides a structured and reproducible framework to support decision-making with greater consistency and transparency.

Despite the progress made, this work has some limitations that open opportunities for future research. The methodology was applied to a single deposit with specific geological characteristics, so it is recommended to validate its performance in other geological contexts and with different sampling configurations. Additionally, although the model considers key geostatistical variables, it does not explicitly incorporate lithological or structural information of the deposit, which could be integrated using multi-source approaches or hybrid models. Lithological information can be incorporated by transforming discrete values into continuous values, which can be achieved by calculating the distance to the boundary of the contiguous lithology. Future research could also explore the use of deep learning techniques to model non-linear variability more efficiently, as well as Bayesian methods to better quantify uncertainty in category assignment. Additionally, spatial clustering techniques such as spatially constrained K-Means or spatial Gaussian mixture models represent a promising alternative to introduce spatial continuity directly into the classification process. These strategies may contribute to a more automated, robust, and geologically consistent framework for mineral resource classification, aligned with international reporting standards.

References

[1]. CIM. (2019). Estimation of mineral resources & mineral reserves best practice guidelines. *Canadian Institute of Mining*, 1-29.

[2]. CRIRSCO. (2024). International reporting template for the public reporting of exploration targets, exploration results, mineral resources and mineral reserves. *International Council on Mining and Metals*, 1-78.

[3]. JORC. (2012). Australasian code for reporting of mineral resources and ore reserves (the JORC Code). *The Australasian Institute of Mining and Metallurgy, Australian Institute of Geoscientists and Minerals Council of Australia*, 1-44.

[4]. CIM. (2014). CIM definition standards for mineral resources & mineral reserves. *CIM Standing Committee on Reserve Definitions*, 1-9.

[5]. SAMREC. (2016). The South African code for the reporting of exploration results, mineral resources and mineral reserves (the SAMREC Code). *The South African Mineral Resource Committee (SAMREC), under the auspices of the South African Institute of Mining and Metallurgy*, 1-96.

[6]. PERC. (2021). Pan-European standard for the public reporting of exploration results, mineral resources and mineral reserves. *Pan-European Reserves and Resources Reporting Committee*, 1-106.

[7]. SME. (2017). SME guide for reporting exploration information, mineral resources, and mineral reserves. *Resources and Reserves Committee of the Society for Mining, Metallurgy, and Exploration*, 1-97.

[8]. CH20235. (2015). Código para informar sobre los resultados de exploración, recursos minerales y reservas minerales. *Comisión Calificadora de Competencias en Recursos y Reservas Mineras (Chile)*, 1-81.

[9]. Li, J., Yang, Q., Liu, Y., & Ma, Y. (2021). Mapping between "Classifications for Mineral Resources and Mineral Reserves (GB/T 17766-2020)" and "United Nations Framework Classification for Resources (UNFC-2019)". *Springer Series in Geomechanics and Geoengineering*, 3405-3416.

[10]. Zuo, M., & Wang, T. (2021). Research on reserve classification of solid mineral resources in China and western countries. *IOP Conference Series: Earth and Environmental Science*, *631*(1), 012044.

[11]. Battalgazy, N., & Madani, N. (2019). Categorization of mineral resources based on different geostatistical simulation algorithms: A case study from an iron ore deposit. *Natural Resources Research*, *28*, 1329-1351.

[12]. Stephenson, P. (2011). Mineral resource classification: How the viability of your project may hang on a qualified person's judgement. *Canadian Institute of Mining*, 1-12.

[13]. Abzalov, M. (2016). Methodology of the mineral resource classification. *Modern Approaches in Solid Earth Sciences*, *12*, 355-363.

[14]. Cevik, I., Leuangthong, O., Caté, A., & Ortiz, J. (2021). On the use of machine learning for mineral resource classification. *Mining, Metallurgy & Exploration*, *38*, 2055-2073.

[15]. Menin, R., Diedrich, C., Reuwsaat, J., & De Paula, W. (2017). Drilling grid analysis for defining open-pit and underground mineral resource classification through production data. In: Gómez-Hernández, J., Rodrigo-Illari, J., Rodrigo-Clavero, M., Cassiraga, E., & Vargas-Guzmán, J. (Eds.). *Geostatistics Valencia 2016*, *19*, 243-254.

[16]. Owusu, S., & Dagdelen, K. (2019). Critical review of mineral resource classification techniques in the gold mining industry. *Insights in Mining Science & Technology*, *1*(3), 555564.

[17]. McManus, S., Rahman, A., Coombes, J., & Horta, A. (2021). Uncertainty assessment of spatial domain models in early stage mining projects – A review. *Ore Geology Reviews*, *133*, 104098.

[18]. de Souza, L., Costa, J., & Koppe, J. (2009). A geostatistical contribution to assess the risk embedded in resource classification methods. *In: Iron Ore 2009 Proceedings*. Perth, WA, 1-12.

[19]. Dohm, C. (2005). Quantifiable mineral resource classification: A logical approach. In: Leuangthong, O., & Deutsch, C. V. (Eds.). *Geostatistics Banff 2004*, *14*, 375-386.

[20]. Sinclair, A., & Blackwell, G. (2000). Resource/reserve classification and the qualified person. *CIM Bulletin*, *93*(1038), 29-35.

[21]. Silva, D., & Boisvert, J. (2014). Mineral resource classification: A comparison of new and existing techniques. *Journal of the Southern African Institute of Mining and Metallurgy*, *114*(3), 265-273.

[22]. Rocha, V., & Bassani, M. (2023). Practical application of a multi-layer scorecard workflow (MLSW) for comprehensive mineral resource classification. *Applied Earth Science: Transactions of the Institute of Mining and Metallurgy*, *132*(3-4), 215-230.

[23]. Stephenson, P., Allman, A., Carville, D., Stoker, P., Mokos, P., Tyrrell, J., & Burrows, T. (2006). Mineral resource classification – It's time to shoot the 'spotted dog'! *Australasian Institute of Mining and Metallurgy Publication Series*, 1-12.

[24]. Isatelle, F., & Rivoirard, J. (2019). Mineral resources classification of a nickel laterite deposit: Comparison between conditional simulations and specific areas. *Journal of the Southern African Institute of Mining and Metallurgy*, *119*(6), 1-9.

[25]. Lepikhina, O., & Pravdina, E. (2016). Comparative analysis of the reserves and estimated resources classification of solid minerals of state committee of reserves and the JORC Code. *International Journal of Applied Engineering Research*, *10*(20), 40950-40955.

[26]. Hernández, H. (2024). A semiautomatic multi-criteria method for mineral resources classification. *Applied Earth Science: Transactions of the Institutions of Mining and Metallurgy*, *133*(2), 211-223.

[27]. Cotrina, M., Marquina, J., & Mamani, J. (2025). Application of artificial neural networks for the categorization of mineral resources in a copper deposit in Peru. *World Journal of Engineering*.

[28]. Cotrina, M., Marquina, J., Mamani, J., Arango, S., Ccatamayo, J., Gonzalez, J., Donaires, T., & Calla, M. (2025). Categorization of mineral resources using Random Forest model in a copper deposit in Peru. *Journal of Mining and Environment*, *16*(4), 947-962.

[29]. Afzal, P., Gholami, H., Madani, N., Yasrebi, A., & Sadeghi, B. (2023). Mineral resource classification using geostatistical and fractal simulation in the Masjed Daghi Cu–Mo porphyry deposit, NW Iran. *Minerals*, *13*(3), 370.

[30]. Madani, N. (2020). Mineral resource classification based on uncertainty measures in geological domains. *In: Springer Series in Geomechanics and Geoengineering*, 157-164.

[31]. Mery, N., Maleki, M., País, G., Molina, A., Cáceres, A., & Emery, X. (2025). Fuzzy classification of mineral resources: Moving toward overlapping categories to account for geological, economic, metallurgical, environmental, and operational criteria. *Natural Resources Research*, *34*, 1271-1299.

[32]. Marquina-Araujo, J., Cotrina-Teatino, M., Cruz-Galvez, J., Noriega-Vidal, E., & Vega-Gonzalez, J. (2024). Application of autoencoders neural network and K-means clustering for the definition of geostatistical estimation domains. *Mathematical Modelling of Engineering Problems*, *11*(5), 1207-1218.

[33]. Sadeghi, M., Casey, P., Carranza, E., & Lynch, E. (2024). Principal components analysis and K-means clustering of till geochemical data: Mapping and targeting of prospective areas for lithium exploration in Västernorrland Region, Sweden. *Ore Geology Reviews*, *167*, 106002.

[34]. Supajaidee, N., Chutsagulprom, N., & Moonchai, S. (2024). An adaptive moving window kriging based on K-means clustering for spatial interpolation. *Algorithms*, *17*(2), 57.

[35]. Pennec, X., Fillard, P., & Ayache, N. (2006). A Riemannian framework for tensor computing. *International Journal of Computer Vision*, *66*(1), 41-66.

[36]. Zhuo, F., Zhang, X., Tang, F., Yu, Y., & Liu, L. (2024). Riemannian transfer learning based on log-Euclidean metric for EEG classification. *Frontiers in Neuroscience*, *18*, 1381572.

[37]. Miah, A., Rahim, M., & Shin, J. (2020). Motor-imagery classification using Riemannian geometry with median absolute deviation. *Electronics (Switzerland)*, *9*(10), 1584.

[38]. Riquelme, Á., & Ortiz, J. (2024). A Riemannian tool for clustering of geo-spatial multivariate data. *Mathematical Geosciences*, *56*, 121-141.

[39]. Bianchi, M., & Zheng, C. (2009). SGeMS: A free and versatile tool for three-dimensional geostatistical applications. *Ground Water*, *47*(1), 8-12.

[40]. Zhang, M., Zhang, Y., & Yu, G. (2017). Applied geostatistics analysis for reservoir characterization based on the SGeMS (Stanford Geostatistical Modeling Software). *Open Journal of Yangtze Oil and Gas*, *2*(1), 45-66.

[41]. Daya, A. (2015). Ordinary kriging for the estimation of vein type copper deposit: A case study of the Chelkureh, Iran. *Journal of Mining and Metallurgy A: Mining, 51*(1), 1-14.

[42]. Shi, N., Liu, X., & Guan, Y. (2010). Research on K-means clustering algorithm: An improved K-means clustering algorithm. *In: Proceedings of the 3rd International Symposium on Intelligent Information Technology and Security Informatics (IITSI 2010)*, China, 1-5.

[43]. Vattani, A. (2011). K-means requires exponentially many iterations even in the plane. *Discrete & Computational Geometry*, *45*(3), 596-616.

[44]. Barachant, A., Bonnet, S., Congedo, M., & Jutten, C. (2013). Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Neurocomputing*, *112*, 172-178.

[45]. Goh, A., & Vidal, R. (2008). Clustering and dimensionality reduction on Riemannian manifolds. *In: Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008)*, USA, 1-8.

[46]. Kalunga, E., Chevallier, S., Barthélemy, Q., Djouani, K., Monacelli, E., & Hamam, Y. (2016). Online SSVEP-based BCI using Riemannian geometry. *Neurocomputing*, *191*, 55-68.

[47]. You, K., & Park, H. J. (2021). Re-visiting Riemannian geometry of symmetric positive definite matrices for the analysis of functional connectivity. *NeuroImage*, *225*, 117464.

[48]. Rodrigues, P., Jutten, C., & Congedo, M. (2019). Riemannian Procrustes analysis: Transfer learning for brain-computer interfaces. *IEEE Transactions on Biomedical Engineering*, *66*(8), 2390-2401.

[49]. Cotrina, M., Marquina, J., Mamani, J., Arango, S., Gonzalez, J., Ccatamayo, J., & Noriega, E. (2024). Predictive model using machine learning to determine fuel consumption in CAT-777F mining equipment. *International Journal of Mining and Mineral Engineering*, *15*(2), 147-160.

[50]. Lei, T., Ng, S., & Siu, S. (2023). Application of ANN, XGBoost, and other ML methods to forecast air quality in Macau. *Sustainability (Switzerland)*, *15*(6), 5341.

[51]. Li, T., Xia, Q., Ouyang, Y., Zeng, R., Liu, Q., & Li, T. (2024). Prospectivity and uncertainty analysis of tungsten polymetallogenic mineral resources in the Nanling Metallogenic Belt, South China: A comparative study of AdaBoost, GBDT, and XGBoost algorithms. *Natural Resources Research*, *33*, 1049-1071.

[52]. Nabavi, Z., Mirzehi, M., Dehghani, H., & Ashtari, P. (2023). A hybrid model for back-break prediction using XGBoost machine learning and metaheuristic algorithms in Chadormalu iron mine. *Journal of Mining and Environment*, *14*(2), 689-712.

[53]. Marquina, J., Cotrina, M., Mamani, J., Noriega, E., & Vega, J., & Cruz, J. (2024). Copper ore grade prediction using machine learning techniques in a copper deposit. *Journal of Mining and Environment*, *15*(3), 1011-1027.

[54]. Farhadi, S., Tatullo, S., Boveiri Konari, M., & Afzal, P. (2024). Evaluating StackingC and ensemble models for enhanced lithological classification in geological mapping. *Journal of Geochemical Exploration*, *260*, 107441.

[55]. Farhadi, S., Afzal, P., Konari, M., Saein, L., & Sadeghi, B. (2022). Combination of machine learning algorithms with concentration-area fractal method for soil geochemical anomaly detection in sediment-hosted Irankuh Pb–Zn deposit, Central Iran. *Minerals*, *12*(6), 689.

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

[57]. Carranza, E., & Laborte, A. (2015). Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). *Computers & Geosciences*, *74*, 60-70.

[58]. Cotrina, M., Marquina, J., Noriega, E., Mamani, J., Ccatamayo, J., Gonzalez, J., & Arango, S. (2024). Predicting open pit mine production using machine

learning techniques: A case study in Peru. *Journal of Mining and Environment*, *15*(4), 1345-1355.

[59]. Ford, A. (2020). Practical implementation of random forest-based mineral potential mapping for porphyry Cu–Au mineralization in the Eastern Lachlan Orogen, NSW, Australia. *Natural Resources Research*, *29*, 267-283.

[60]. Lachaud, A., Adam, M., & Mišković, I. (2023). Comparative study of random forest and support vector machine algorithms in mineral prospectivity mapping with limited training data. *Minerals*, *13*(8), 1073.

[61]. Chaudhry, M., Shafi, I., Mahnoor, M., Vargas, D., Thompson, E., & Ashraf, I. (2023). A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective. *Symmetry (Basel)*, *15*(9), 1679.

[62]. Tarigan, D. (2023). Optimization of the K-means clustering algorithm using Davies–Bouldin index in Iris data classification. *KLIK: Media Informasi dan Komunikasi Penelitian*, *4*(1), 1-7.

انجمن مهندسی معدن ایران
نشریه مهندسی معدن و محیط زیست
نشانی نشریه: www.jme.shahroodut.ac.ir
دانشگاه صنعتی شاهرود

یک رویکرد نوآورانه برای طبقه‌بندی منابع معدنی بر اساس خوشه‌بندی ریمانی و یادگیری ماشین در یک کانسار مس

مارکو آنتونیو کوترینا-تئاتینو[1,4*]، جایرو جوناتان مارکوینا-آرائوخو[1,4]، خوزه نستور مامانی-کوئیسپ[2,4]، سولیو مارینو آرانگو-رتاموزو[1,4]، جو الکسیس گونزالس-واسکوئز[3] و کوین دانیل روندو-جالکا[1,4]

۱ .گروه مهندسی معدن، دانشکده مهندسی، دانشگاه ملی تروخیو، تروخیو، پرو

۲ .دانشکده مهندسی شیمی، دانشگاه ملی آلتیپلانو پونو، پونو، پرو

۳ .گروه مهندسی صنایع، دانشکده مهندسی، دانشگاه ملی تروخیو، تروخیو، پرو

۴. آزمایشگاه طراحی و برنامه‌ریزی معدن – DIPLAMIN، دانشگاه ملی تروخیو، تروخیو، پرو

اطلاعات مقاله

تاریخ ارسال: ۲۰۲۵/۰۵/۲۸

تاریخ داوری: ۲۰۲۵/۰۶/۲۷

تاریخ پذیرش: ۲۰۲۵/۰۹/۰۶

DOI: 10.22044/jme.2025.16320.3174

کلمات کلیدی

طبقه‌بندی منابع معدنی

هندسه ریمانی

زمین‌آمار

چکیده

طبقه‌بندی منابع معدنی به طور قابل توجهی بر برنامه‌ریزی معدن، امکان‌سنجی اقتصادی و رعایت مقررات تأثیر می‌گذارد. با وجود اهمیت آن، چنین طبقه‌بندی اغلب به قضاوت ذهنی فرد واجد شرایط (QP) بستگی دارد، زیرا معیارهای فنی استاندارد بین‌المللی برای تعیین دسته‌بندی منابع وجود ندارد. برای کاهش این محدودیت، یک روش نوآورانه که خوشه‌بندی مبتنی بر هندسه ریمانی را با تکنیک‌های یادگیری ماشین برای طبقه‌بندی منابع معدنی ادغام می‌کند، توسعه داده شد. یک پایگاه داده از ۵۶۵۴ نمونه ترکیبی از ۱۸۵ گمانه مته الماس در یک کانسار مس در مرکز پرو برای طبقه‌بندی ۳۱۸۴۴۳ بلوک استفاده شد. عیار مس از طریق کریجینگ معمولی (RMSE = 0.102؛ MAE = 0.069) تخمین زده شد و متغیرهای زمین‌آماری واریانس کریجینگ، میانگین فاصله تا نمونه‌ها و تعداد نمونه‌هایی که به عنوان ویژگی‌های ورودی برای طبقه‌بندی عمل می‌کردند، تولید شدند. خوشه‌بندی با استفاده از KMeans کلاسیک و KMeans ریمانی انجام شد و پس از آن هموارسازی مکانی از طریق الگوریتم‌های XGBoost و جنگل تصادفی انجام شد. مختصات مطلق برای رسیدگی به ناپیوستگی‌های مکانی در خروجی‌های طبقه‌بندی گنجانده شدند. ترکیب مدل ریمانی با جنگل تصادفی، بالاترین عملکرد طبقه‌بندی را با شاخص سیلوئت ۰.۲۶ و شاخص دیویس-بولدین ۰.۷۲ به همراه داشت. محتوای فلزی حاصل، ۴.۲۴ میلیون تن مس با عیار ۰.۴۴٪ (اندازه‌گیری شده)، ۶.۴۹ میلیون تن با عیار ۰.۳۴٪ مس (نشان داده شده) و ۷.۶۸ میلیون تن با عیار ۰.۳۲٪ مس (استنباطی) تخمین زده شد که نشان‌دهنده همسویی نزدیک با تخمین‌های QP و در عین حال بهبود انسجام مکانی است. به طور خلاصه، رویکرد مبتنی بر ریمانی نسبت به روش‌های طبقه‌بندی KMeans کلاسیک و مرسوم عملکرد بهتری داشت و جایگزینی قوی‌تر، عینی‌تر و سازگارتر در سطح جهانی ارائه داد.

نویسنده مسئول مکاتبات: *mcotrinat@unitru.edu.pe* (م. کوترینا-تئاتینو)