Shahrood University of Technology

Journal of Mining and Environment (JME)

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

Iranian Society of Mining Engineering (IRSME)

Comparative Analysis of Machine Learning Regression Methods for Geometallurgical Modeling in the Sungun Copper Porphyry Deposit

Meysam Nikfarjam[1], Ardeshir Hezarkhani[1*], Farhad Azizafshari[2], and Hamidreza Golchin[2]

1. Faculty of Mining Engineering, Amirkabir University of Technology, Tehran, Iran.
2. National Iranian Copper Industries Co. (NICICO), Sungun Copper Mine, East-Azerbaija, Iran.

Article Info

Received 18 September 2024
Received in Revised form 27 March 2025
Accepted 16 May 2025
Published online 16 May 2025

DOI: 10.22044/jme.2025.15088.2885

Keywords

Geometallurgical modeling
Process properties
Semi-autogenous Power Index (SPI)
Machine Learning Algorithms
Sungun Copper mine

Abstract

Geometallurgical modeling (GM) plays a crucial role in the mining industry, enabling a comprehensive understanding of the complex relationship between geological and metallurgical factors. This study focuses on evaluating metallurgical varibles at the Sungun Copper mine in Iran by measuring and predicting process properties, including semi-autogenous power index (SPI), recovery (Re), and concentration grade. To overcome the additivity limitations of geostatistical methods, we utilized machine learning algorithms for enhanced predictive modeling, aiming to improve decision-making and optimize mining operations in geometallurgy. The research incorporates crucial data inputs such as sample coordinates, grades, lithology, mineralization zones, and alteration to assess the accuracy and reliability of different machine learning regression methods. The Relative Standard Deviation (RSD) is highlighted as a significant metric for comparing the accuracy of predicted process properties. Evaluation metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and coefficient of determination (R2) further confirm the superiority of specific modeling methods in certain scenarios. The K-Nearest Neighbors (KNN) method exhibits superior accuracy, lower error metrics (RMSE and MAE), and a higher R2 for modeling the SPI test. For modeling Cu grade in concentrate, Support Vector Regression (SVR) proves to be effective and reliable, outperforming the Multilayer Perceptron (MLP) method. Despite MLP's high R2, its higher RSD suggests increased uncertainty and variability in the predictions. Therefore, SVR is considered more suitable for modeling Cu grade in concentrate. Findings optimize operations at Sungun Copper mine, improving decision-making, efficiency, and profitability.

1. Introduction

Geometallurgical modeling (GM) is a comprehensive approach that integrates data from various disciplines, including geology, metallurgy, and mineral processing to optimize mining operations [1-6]. GM encompasses the comprehensive characterization of ore deposits, integrating both geological and metallurgical aspects. This approach enables more precise predictions of process performance and resource recovery while optimizing efficiency and maximizing resource utilization [4]. GM's significance lies in bridging the gap between geology and mineral processing, acknowledging how variations in ore composition, physical properties, and other factors greatly influence overall profitability in mining operations [1, 6, 7]. Geometallurgy enables proactive decision-making by quantifying these relationships through predictive models [4]. In the context of geometallurgy, there are indeed three commonly recognized approaches: traditional, proxy-based, and mineralogical. The traditional approach in geometallurgy integrates geological and metallurgical data separately. Geological data, such as lithology, structure, and alteration, are analyzed to understand the spatial variability of the deposit.

* Corresponding author: ardehez@aut.ac.ir (A. Hezarkhani)

Metallurgical data, including grindability and flotation response, are examined to assess the processability of the ore. With the goal of optimizing operations, the traditional approach in mineral processing takes into account both geological and metallurgical factors [3, 8]. The proxy-based approach in geometallurgy involves using indirect measurements or "proxies" to estimate or predict metallurgical properties. Proxies are easily measurable parameters that are correlated with the target metallurgical properties of interest. This approach is useful when direct measurements are challenging or time-consuming [3, 9, 10]. The mineralogical approach in geometallurgy focuses on the detailed analysis of mineralogical characteristics and their impact on extractive processes. Mineralogical studies involve mineral identification, mineral associations, and liberation analysis. The mineralogical approach provides valuable insights into the behavior of minerals during processing, but it can be costly, resulting in limited data availability, and faces challenges related to sample representativeness and interpretation complexity. [11, 12]. Models that make use of both grades and mineralogy exhibit remarkable prediction performance, while models relying only on grades and geological characterization can still achieve satisfactory results in the absence of mineralogy data. [13-16]. GM involves the use of qualitative and quantitative properties as proxies for metallurgical responses, which are applicable across mentioned approaches. To predict metallurgical responses, regression models are fitted using primary variables within this framework. The lack of metallurgical response property data makes it difficult to establish predictive response relationships, even though primary rock property data is relatively abundant. Complex relationships between primary input variables and geometallurgical responses add to the complexity of the prediction process. For prediction purposes, it is necessary to find a better alternative to traditional multivariate linear regression models that perform poorly in primary-response relationships [4].

Machine-learning (ML), a discipline within the field of artificial intelligence (AI), concentrates on exploring mathematical models and algorithms to generate predictions. These predictions are derived by analyzing patterns and data, connecting them with existing knowledge, and developing learning algorithms [17-19]. In the realm of geosciences, ML methods bring forth several advantages over alternative approaches, posing a significant challenge to established geostatistical and Bayesian methods [20]. While ML algorithms are extensively employed in geoscience applications such as geochemical anomaly detection [21, 22], lithological classification for geological mapping [23], separation of alteration zones [24] and mineral prospectivity mapping [25-27], their usage in GM is relatively uncommon. However, over the past decade, some researchers have started utilizing ML in the geometallurgical domaining [28] and geometallurgical 3D modeling [29]. ML methods possess a significant advantage in their capability to effectively handle the additivity issue. Additivity, as a mathematical property, enables the computation of variable means through linear averaging. Comminution and recovery are examples of geometallurgical response variables that exhibit non-additivity [29-31]. Geometallurgists can enhance their predictive modeling capabilities by employing advanced statistical techniques, including ML algorithms such as support vector machines (SVM), random forests (RF), or neural networks (NN). By utilizing these methods, robust models can be developed to estimate critical parameters required for efficient mine planning and optimization. This enables more accurate and effective decision-making in the field of geometallurgy [1, 4, 29].

Several tests are commonly employed to assess the performance of comminution processes. These include the Bond mill work index (BWi), Bond rod mill work index (RWi), SAG Power Index (SPI), drop-weight index (DWi), and resistance to abrasion and breakage index (A×b). The primary methods used to assess processing recovery are typically centered around either flotation or leaching operations. These tests serve as common approaches to evaluate the efficiency and effectiveness of recovery processes [14, 32].

In this particular study, the SPI test was selected to investigate the variability in semi-autogenous hardness. The test was conducted using the Starkey laboratory mill commercialized by Minnovex, which had specific dimensions and equipment for the purpose. The SPI tests involved grinding 2 kg samples to specific particle sizes, with the results used to determine the time required to achieve a particular size distribution. A higher SPI value indicates increased resistance to grinding, indicating a harder ore [33]. The SAG mill specific energy equation, developed by [34], links mill power draw per unit throughput to the SPI and T_{80}parameters as follows:

$$E_{SAG}(kWh/t) = \frac{2.2 + 0.1SPI}{T_{80}^{0.33}} \quad (1)$$

where E_{SAG} represents the mill specific energy (kWh/t), $T_{80}(mm)$ is the 80% passing size of the SAG screen undersize, and SPI(min), denotes the time required to achieve this specific size distribution.

Flotation recovery and concentrate grade were another geometallurgical variable utilized in this study. Flotation is a highly effective mineral separation method widely used in mineral processing for separating valuable mineral particles from gangues based on their different physicochemical properties. It is commonly applied in processing copper sulfide ores. The effectiveness of the flotation process is influenced by both feed characteristics and operational parameters. The flotation characteristics of copper oxide minerals vary from those of copper sulfide minerals, leading to reduced recovery rates in flotation procedures [35]. Flotation tests were carried out by preparing a pulp using 471 grams of sample in a 1.3L cell, with specific collectors and frothers added. After adjusting the pH to 11, the flotation tests were conducted following the mixing of the pulp and chemicals. Following that, air was introduced into the flotation cell, and the concentrate was collected. In each test, the recovery of the concentrate is calculated using the following Equation [36]:

$$Recovery, Re(\%) = \frac{c(f-t)}{f(c-t)} \times 100 \quad (2)$$

Where '*c*' denotes the concentrate grade, '*f*' and '*t*' represent the feed and tail grade, respectively.

Copper porphyry mines are known for their complex ore characteristics and variability, posing challenges to accurate GM. In this paper, due to certain limitations in accessing the boreholes for validation, we perform a comparative analysis of various ML regression methods. This involves evaluating their performance using error metrics such as relative standard deviation (RSD), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and other pertinent measures. Nevertheless, this study can serve as a valuable reference for selecting an appropriate machine-learning regression method to model process data of a porphyry copper deposit. The study employs samples obtained from the Sungun Porphyry copper deposit. It focuses on evaluating the integration of three key process properties, namely SPI, Copper recovery and Copper grade in concentrate, into a spatial model using machine-learning methods. The primary objective of our evaluation of different algorithms is to pinpoint the regression method that is most appropriate for accurately predicting significant metallurgical parameters. This selection will ultimately contribute to advancements in resource estimation, process optimization, and overall operational efficiency.

2. Methods

In GM, three primary challenges often arise. The first challenge is the limited number of geometallurgical samples compared to geological and grade variables, which hinders robust modeling due to the high cost of sampling. The second challenge is the prevalence of non-additivity among geometallurgical variables, such as recovery, which is commonly overlooked when using traditional estimation methods like kriging, leading to biased results in standard geostatistical estimations or simulations [4, 31, 37]. Instead of relying solely on traditional estimation methods like kriging, incorporating advanced modeling techniques that can handle non-additive variables can help mitigate bias in estimations. This may include the use of ML algorithms, geostatistical simulation methods, or other non-linear regression approaches. Based on the current information, ML has not been widely employed in the spatial modeling of process properties. Instead, the dominant techniques used thus far have been regression models, geostatistical, and multivariate statistics methods [32, 38]. In this study, various ML regression algorithms were utilized to determine the most appropriate method for geometallurgical modeling. By employing a range of algorithms, the main advantage is the ability to compare and contrast their performance in predicting and modeling complex geological and metallurgical data. Each algorithm has unique strengths and weaknesses, allowing for a comprehensive evaluation of their predictive capabilities and aiding in the selection of the most effective and accurate method for the specific characteristics of the dataset.

The data processing for the application of ML algorithms (MLA) involved three primary stages. Firstly, the algorithms were trained and parameterized to optimize their performance. Secondly, post-processing was conducted, which entailed converting the output values into a map format. Lastly, an accuracy assessment was performed to evaluate the performance of the MLA models. It is worth noting that all of these MLA models were developed using the Python programming language, highlighting its

significance as the chosen platform for implementing the models. To thoroughly examine the performance of various ML algorithms, it is crucial to identify appropriate parameters for each model. By carefully selecting and fine-tuning parameters, researchers can create effective and reliable predictive models optimized for their specific application and dataset. The way in which MLAs are parameterized significantly affects their robustness and ability to generalize, thereby impacting their accuracy in predicting new response variables.

2.1. Support Vector Regression (SVR)

In regression tasks, Support Vector Regression (SVR) serves as a powerful ML algorithm capable of effectively capturing the intricate and nonlinear relationships that exist between predictor variables and response variables. SVR builds upon the principles of Support Vector Machines (SVM) and aims to find a hyperplane that best fits the data while minimizing deviations within a specified margin [39, 40]. SVR is widely acknowledged as the preferred approach for regression tasks with limited sample sizes. It is well-suited for modeling various types of non-linear relationships by utilizing a kernel function to construct functions based on representative samples, known as support vectors, selected from the entire sample set [41].

To solve the optimization problem in SVR, the formulation includes the introduction of slack variablesξ and $\hat{\xi}$which accommodate errors or deviations from the margin.

To analyze each data point, two slack variables $\xi_n \geq 0$ and $\hat{\xi}_n \geq 0$ are required. Specifically, $\xi_n > 0$ represents a point where $t_n > y(x_n) + \varepsilon$, while $\hat{\xi}_n > 0$ represents a point where $t_n < y(x_n) - \varepsilon$. This is visually depicted in Figure 1. The objective function aims to minimize the complexity of the model and the error terms, while adhering to predefined constraints.

By solving this optimization problem using various techniques like quadratic programming or gradient descent, SVR finds an optimal hyperplane that maximizes prediction accuracy while minimizing errors within a defined tolerance zone around it [39, 40].

2.2. Random Forest Regression

Random Forest (RF) Regression is a powerful ML algorithm that utilizes an ensemble of decision trees to perform regression tasks. It combines the predictions of multiple individual decision trees to generate a more accurate and robust prediction [40]. The basic idea behind RF regression can be summarized as follows [39, 40, 42].

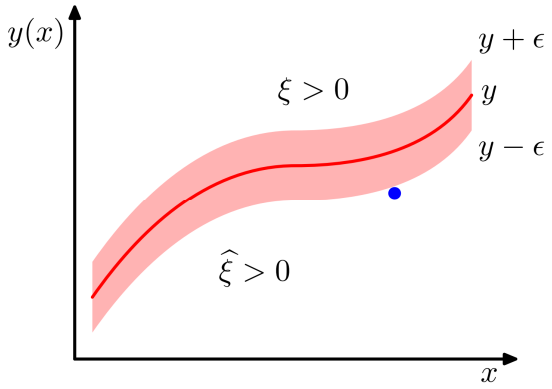


Figure 1. Visualization of SVM regression, depicting the regression curve alongside with the ε −insensitive 'tube' [39].

Building Decision Trees: A large number of decision trees are constructed using random subsets of the training data.

Bootstrap Aggregation (Bagging): Each tree is trained on a different subset of the training data obtained through bootstrapping, which involves sampling from the original dataset with replacement.

Feature Subsetting: At each node split within each tree, only a subset of features (randomly selected) are considered for determining the best split, reducing correlation between trees.

Predicting by Averaging: The predictions from all individual trees are averaged or combined using another technique such as weighted averaging or majority voting to obtain the final prediction.

The general formula for predicting using RF regression involves aggregating predictions from multiple decision trees:

Prediction = Average (Prediction_Tree_1, Prediction_Tree_2,... , Prediction_Tree_n)

Where Prediction_Tree_i represents the prediction made by Tree i.

This averaging process helps reduce overfitting and improves model stability by considering diverse perspectives captured by different decision trees within the forest.

2.3. MLP

Regarded as a highly successful model for pattern recognition, the multilayer perceptron (MLP), commonly referred to as a feed-forward neural network, has established its prominence in the field. Interestingly, the term "multilayer perceptron" can be somewhat misleading. This is

because the model is actually comprised of multiple layers of logistic regression models, which exhibit continuous nonlinearities, rather than multiple perceptrons with discontinuous nonlinearities. In many cases, this leads to a more concise model that can be evaluated faster than a SVM with comparable generalization performance. However, achieving this compactness comes at a cost. Similar to the SVM, the likelihood function used for network training is no longer a convex function of the model parameters. Nevertheless, investing significant computational resources during the training phase to obtain a compact model that efficiently processes new data is often worthwhile in practice [39].

The neural network model shown in Figure 2 features two processing stages similar to the perceptron model, leading to its designation as a multilayer perceptron (MLP). A key difference is that the neural network uses continuous sigmoidal nonlinearities in its hidden units, while the perceptron relies on step-function nonlinearities. This difference is crucial as it allows the neural network to be differentiable with respect to its parameters, which is essential for effective training.

The structure of Artificial Neural Networks (ANNs), as depicted in Figure 2, comprises three essential elements: an input layer, a hidden layer, and an output layer. All the concepts can be precisely expressed using mathematical terms, as described by Equation (3).

$$z_j = h(a_j), \qquad a_j = \sum_{i=1}^{D} w_{ji}{}^{(1)} x_i + w_{j0}{}^{(1)} \tag{3}$$

where x_1, ..., x_D are input variables, j ranges from 1 to M (number of linear combinations), and the superscript (1) denotes the parameters in the first 'layer' of the network, the parameters $w_{ji}{}^{(1)}$ are referred to as weights, and the parameters $w_{j0}{}^{(1)}$ are known as biases. The quantities a_j are called activations. To obtain the values of z, each activation a_j is subjected to a differentiable, nonlinear activation function $h(\cdot)$.

Sigmoidal functions, like the *logistic sigmoid* function, are commonly chosen as the nonlinear functions $h(\cdot)$. Building upon equation (3), these values are further combined in a linear manner to produce the activations of the output units.

$$a_k = \sum_{j=1}^{M} w_{kj}{}^{(2)} z_j + w_{k0}{}^{(2)} \tag{4}$$

The transformation aligns with the second layer of the network, where k ranges from 1 to K, with K indicating the total number of outputs. The parameters $w_{k0}{}^{(2)}$ are once again bias parameters in this layer. Subsequently, the output unit activations undergo another transformation using a suitable activation function, resulting in a set of network outputs denoted as y_k. Similar to linear models, the selection of activation function in this context is guided by the characteristics of the data and the distribution of target variables. The activation of each output unit is transformed using a logistic sigmoid function, resulting in the output value denoted as

$$y_k = \sigma(a_k) \quad , \quad \sigma(a) = \frac{1}{1 + exp(-a)} \tag{5}$$

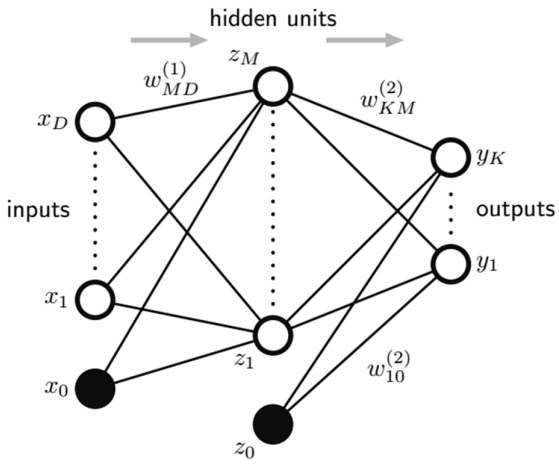


Figure 2. The network diagram of the two-layer neural network illustrates the representation of input, hidden, and output variables as nodes. The weight parameters are illustrated in the diagram as connections between the nodes, whereas the bias parameters are depicted as links extending from additional input and hidden variables, such as x0 and z0. The arrows in the diagram signify the direction of information flow during the process of forward propagation through the network [39].

2.4. K-Nearest Neighbors (KNN)

KNN regression is a highly effective technique for regression tasks, particularly when the data distribution is unknown or complex. It leverages the concept of considering the K nearest neighbours to capture local patterns and relationships within the data, resulting in accurate predictions. The lazy learning nature of KNN regression enables it to swiftly adapt to new data and accommodate any modifications in the

underlying target function [43]. In KNN regression, when given an input x from the training data, it considers K observations with similar xi values and calculates the average response of those K data points. This average response, denoted as $\hat{y}$, is used to estimate the target output for the given input x, as shown in Equation (6):

$$\hat{y}(x) = \frac{1}{k} \sum_{x_i \in N_k(x)} y_i \quad (6)$$

The notation $N_k(x)$ represents the K closest points in the neighborhood of x. To determine the closeness between points, various distance measures can be employed, but the Euclidean distance is commonly used in practice [40].

3. Studied Area

The Sungun porphyry copper deposit is located in Iran's East Azarbaijan province, situated in the north-western region of the country (Figure 3). The deposit comprises a sequence of Cretaceous limestone and shale interlayers as the oldest exposed rocks. Overlying the sedimentary rocks in the Sungun deposit are Upper Eocene volcanic breccia and sandstone, which are intruded by a diorite/granodioritic to quartz-monzonitic stock [44]. Within the Sungun deposit, the main porphyry stock consisting of quartz monzonite to granodiorite and granite drives hydrothermal activities and the formation of different alteration zones, including potassic, propylitic, phyllic, and argillic alterations [45-47]. The occurrence of skarn mineralization predominantly at the eastern and northern boundaries of the stock is a result of the metamorphic processes that have occurred [48]. To develop a 3D geometallurgical model in the Sungun deposit, a specific part with higher mineralization potential was selected based on Nikfarjam's study on the geological domains of the deposit [49]. This model integrates various data sources, including assay data, lithology, mineralization zones, SAG power index (SPI), and flotation recovery (Re) of rougher, all measured on selected core sample intervals.

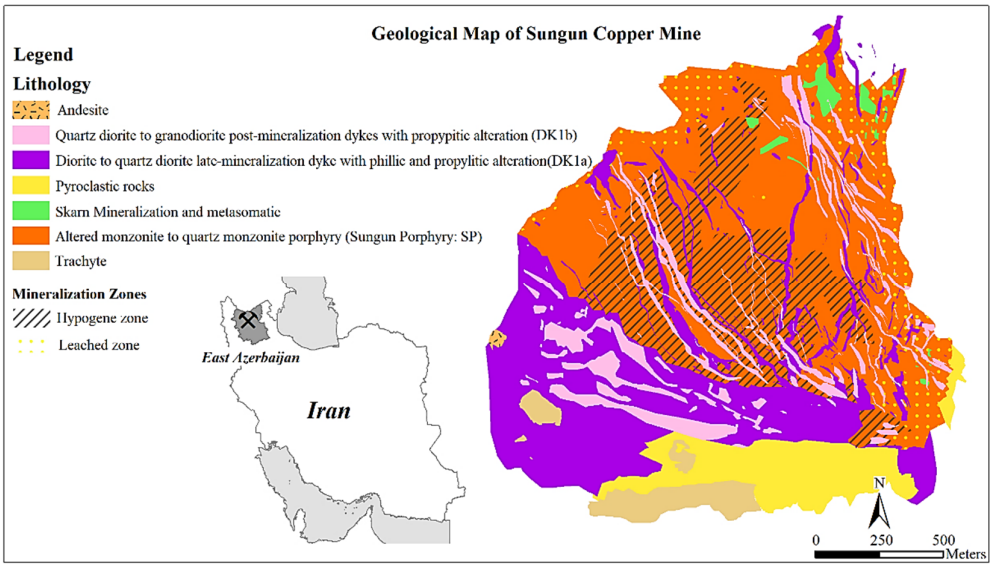


Figure 3. Location map and geological map of Sungun deposit (modified after NICICO).

4. Dataset and Methodology

The study utilized an initial dataset of 45 ore samples extracted from boreholes at 2-meter intervals (each sample about 3 Kg) (Figure 4). These samples were intentionally chosen from the most mineral-rich sections within the ore body's geological domains. By focusing on the highest-grade areas, the dataset provides a targeted representation, allowing for a more accurate assessment and modeling of the ore body's geometallurgical characteristics. The samples were spaced approximately 50 to 100 meters apart, ensuring a spatially distributed dataset (Figure 5). At the metallurgical laboratory of the Sungun copper mine, each ore sample, referred to as "feed," underwent comprehensive geometallurgical tests such as the SAG Power Index (SPI) and rougher froth flotation. Furthermore, all samples underwent analysis using the atomic absorption method with a Shimatsu 7000 machine to detect Cu, CuO, Fe, and Mo. The (Table 3) illustrates the basic statistics

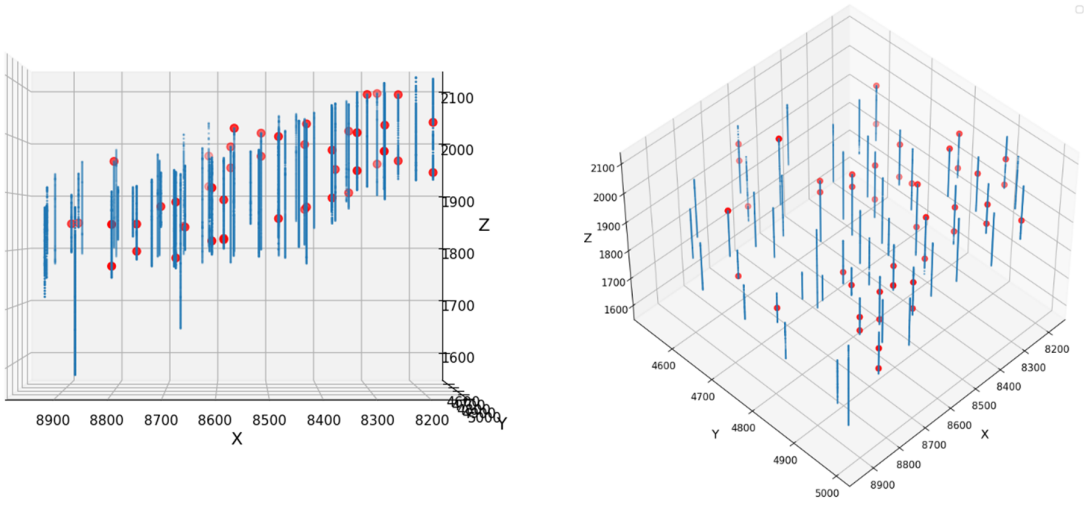
of geometallurgical test results categorized by lithology, alteration, and mineralization. The second dataset used in the study consists of 13,400 samples containing geological and assay data, which can be easily linked to the geometallurgical response dataset. By utilizing primary data that includes quantitative variables (coordinates and assay) and qualitative variables (lithology, alteration, etc.) (Table 2), machine learning (ML) models are employed to estimate the response variables, namely SPI, Cu Recovery, and Cu grade in concentrate in the second dataset, resulting in the creation of a new updated database. The updated database contains coordinates, assay data, geological information, and geometallurgical responses, which can be used to populate the geometallurgical block model. The block model used in this study was created and approved by mining experts for operational mining activities (main block size: 15m*15m*12.5m, total number of blocks: 854453). During the generation of the geometallurgical model, the existing model was enriched with geometallurgical variables and machine learning algorithms. To account for the varying scales of quantitative data and the inclusion of qualitative variables, the initial phase of the study involves standardizing the data and encoding (label encoding) the qualitative variables. This important step ensures that all data points are transformed to a common scale and that qualitative variables are adequately represented. By doing so, accurate comparisons and analyses can be conducted in subsequent stages of the study or modeling process, enabling a more comprehensive understanding of the data.

It is important to note that ore samples are of small volume, with each sample weighing approximately 3 kg and used in testwork. This limited volume leads to a restricted support size in spatial planning. When geometallurgical properties are modeled based on the primary data (coordinates, geological, and assay data) of the feed and integrated into the updated geological database, the support of these samples may increase if their composite length is adequate. The support of blocks in the block model will increase as they have a larger volume compared to composites from the geological database. The spatial model can be employed to populate spatial objects (geological database or block model) with process parameters, assuming spatial correlation of the process properties. In the process of modeling geometallurgical variables, geometallurgical tests are carried out at the laboratory scale and on the scale of ore samples. Numerous geometallurgical variables are determined through laboratory tests conducted on various scales. It may be inaccurate to assume that these variables will scale up linearly from small laboratory tests to larger plant scales. For example, SPI measured in laboratory tests on core samples using specific instrumentation and standardized methods could vary significantly from those measured in plant conditions. Due to the intricate nature of the upscaling issue, which could be a standalone topic for a PhD thesis [4], this research assumes that response variables will scale up from the sample scale to the block scale.

Figure 4. Core Samples from selected boreholes.

Table 1. Description of data types of primary data.

Primary variable *45 ore samples*	Data Type	Including	Significance
Sample Coordinates	Quantitative	X, Y, Z (m)	Facilitates precise spatial data representation and identification of patterns.
Grades	Quantitative	Cu (%),CuO (%), Fe (%), Mo (ppm)	Assesses the ore's quality and potential profitability.
Lithology	Qualitative	SP, DK1a, DK1b	Aids in comprehending the geological context and its influence on metallurgical performance.
Mineralization zones	Qualitative	Hypogene, Supergene-Hypogene, DK	Facilitates the identification of high-grade areas for mining purposes and optimization of resource utilization.
Alteration	Qualitative	Phyllic, Potassic, Argillic, Sericitic	Assists in assessing the influence of alteration on metallurgical performance.



(a) **(b)**

Figure 5. Location map of data: a) XZ plan and b) 3D view of samples and boreholes.

The performance of regression models can be greatly affected by the inclusion or exclusion of individual data points, especially when working with a limited dataset. To mitigate this challenge, a widely adopted approach is k-fold cross validation. In this technique, the dataset is divided into k subsets of equal size, with one subset used for testing and the remaining k-1 subsets utilized for training. This process is repeated k times, ensuring that each subset is used once for validation. By employing k-fold cross validation, the models' performance can be more accurately evaluated, enhancing their ability to generalize to unseen data [16, 29]. In this study, the modeling results are assessed using 10-fold cross-validation with the utilization of the Scikit-learn package.

The objective of this study is to assess and select the most suitable ML regression models by evaluating the estimation of geometallurgical variables in a porphyry copper deposit using both the relative standard deviation (RSD) and the evaluation metrics:

$$relative\ standard\ deviation = RSD = \frac{\sigma}{\mu} \times 100 \tag{7}$$

$$coefficient\ of\ determination = R^2(y, \hat{y}) = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}, \tag{8}$$

$$root mean - squared\ error = RMSE(y, \hat{y}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{9}$$

$$mean\ absolute\ error = MAE(y, \hat{y}) = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{10}$$

$\sigma$is the standard deviation, μ is mean of population, $\hat{y}$is the prediction of $y$and $\bar{y}$is the mean value of y.

Table 2. Statistical information for geometallurgical ore samples and borehole samples both for numerical and categorical data.

Numerical data							
	X (m)	Y (m)	Z (m)	Cu (%)	CuO (%)	Fe (%)	Mo (ppm)
13400 borehole samples							
min	7955.85	4319.48	1552.68	0.0002	0.0001	0.2504	0.07
max	9260.91	5694.25	2173.55	8.9382	2.1018	22.5965	2737.70
median	8607.43	4951.92	1906.56	0.3481	0.0071	3.3066	31.00
average	8608.78	4976.21	1906.27	0.4619	0.0293	3.7433	107.10
St.d.	253.15	331.33	95.39	0.4604	0.0947	2.0371	176.95
count	13400	13400	13400	13400	13400	13400	13400
45 ore samples							
min	8185.58	4536.22	1773.44	0.0037	0.0001	1.3427	0.68
max	8883.78	4997.37	2088.30	1.5818	0.0606	5.2471	816.90
median	8565.61	4842.12	1944.31	0.7016	0.0092	2.9394	131.40
average	8519.07	4800.32	1932.08	0.6740	0.0143	3.0564	190.02
St.d.	194.87	128.34	82.05	0.3813	0.0144	0.9388	192.60
count	45	45	45	45	45	45	45

Number of samples in categorical data						
	Zone	count	Lithology	count	Alteration	count
13400 borehole samples						
	Hypogene	6720	SP	8502	Phyllic	9984
	Supergene	1481	DK1a	2617	Potassic	1627
	DK	1886	DK1b	542	Argillic	66
					Sericitic	145
45 ore samples						
	Hypogene	28	SP	40	Phyllic	30
	Supergene	9	DK1a	4	Potassic	12
	DK	8	DK1b	1	Argillic	2
					Sericitic	1

Table 3. Basic statistics of geometallurgical tests results.

Features			E(kWh/t)			Cu_Recovery (%)			Cu grade in concentrate (%)		
Zone	***Lithology***	***Alteration***	***Min***	***Max***	***Average***	***Min***	***Max***	***Average***	***Min***	***Max***	***Average***
Hypogene			4.38	57.51	29.02	65.36	97.37	92.75	1.25	7.19	3.17
	SP		4.38	57.51	29.02	65.36	97.37	92.75	1.25	7.19	3.17
		Argillic	4.38	4.38	4.38	65.36	65.36	65.36	1.85	1.85	1.85
		Phyllic	15.34	57.51	30.83	86.45	97.14	92.66	1.25	7.19	3.46
		Potassic	25.83	32.17	28.94	96.10	97.37	96.84	1.70	4.60	2.76
Supergene			27.19	58.87	44.09	91.34	95.05	93.48	1.05	6.07	3.34
	SP		27.19	58.87	44.09	91.34	95.05	93.48	1.05	6.07	3.34
		Phyllic	27.19	58.87	44.09	91.34	95.05	93.48	1.05	6.07	3.34
DK			15.14	50.73	28.39	23.22	94.54	75.20	0.08	5.60	1.66
	DK1a		15.14	32.10	22.81	80.53	94.54	88.19	0.30	5.60	2.05
		Phyllic	50.73	32.10	22.81	80.53	94.54	88.19	0.30	5.60	2.05
	DK1b		50.73	50.73	50.73	0.13	0.13	0.13	23.22	23.22	23.22
		Sericitic		50.73	50.73	0.13	0.13	0.13	23.22	23.22	23.22

5. Results and Discussion

These parameterized distinct methods were employed to construct geometallurgical models aimed at populating the geological database and block model with process properties (Table 4 and Figure 6-8). When applying spatial process models to the geometallurgical block model, a limitation arises in estimating process properties compared to drill cores in the geological database. ML methods used in block models may have higher uncertainty because block model grades are estimated from

geostatistical models, while geological database grades are directly measured. Another contributing factor to the difference between spatial process models for block models and the geological database is the variation in support size. Increasing the support size, such as the composite or block size, reduces variance and introduces spatial smoothing. Consequently, although the average values of process properties for drill core samples and mining blocks may be similar, the variability is smaller for larger supports. These limitations highlight the importance of considering higher uncertainty and spatial smoothing effects when using ML methods to estimate process properties in block models. It is crucial to acknowledge these limitations and understand their impact on the accuracy and variability of predicted process properties [29, 50].

In order to evaluate the accuracy and overall predictive performance of the MLA models, a comparison plot was created. The outcomes of the evaluation process are depicted in Figure 9, where each plot corresponds to the modeling of a distinct process property. The vertical axis of the plot represents the RSD expressed in percentage (%), RMSE, and MAE. On the horizontal axis, the different ML methods used for modeling the geometallurgical variables are listed. This graphical representation allows for a comprehensive comparison of the RSD, RMSE, and MAE values across various ML methods, providing valuable insights into the performance of each method in modeling different geometallurgical variables.

Table 4. Summary of statistics of different MLAs for response variables.

Response Variable	MLP			KNN			RF			SVR		
	Mean	*Median*	*Standard deviation*	*Mean*	*Median*	*Standard deviation*	*Mean*	*Median*	*Standard deviation*	*Mean*	*Median*	*Standard deviation*
SPI	37.22	37.20	4.00	34.51	33.60	4.24	32.96	29.80	7.12	26.05	16.57	12.04
Re	92.28	92.21	2.56	93.63	93.86	1.63	91.62	90.23	2.22	94.91	95.44	.92
Cu Concentrate	2.42	2.45	.48	2.90	2.83	.23	2.89	2.99	.15	3.07	3.18	.56

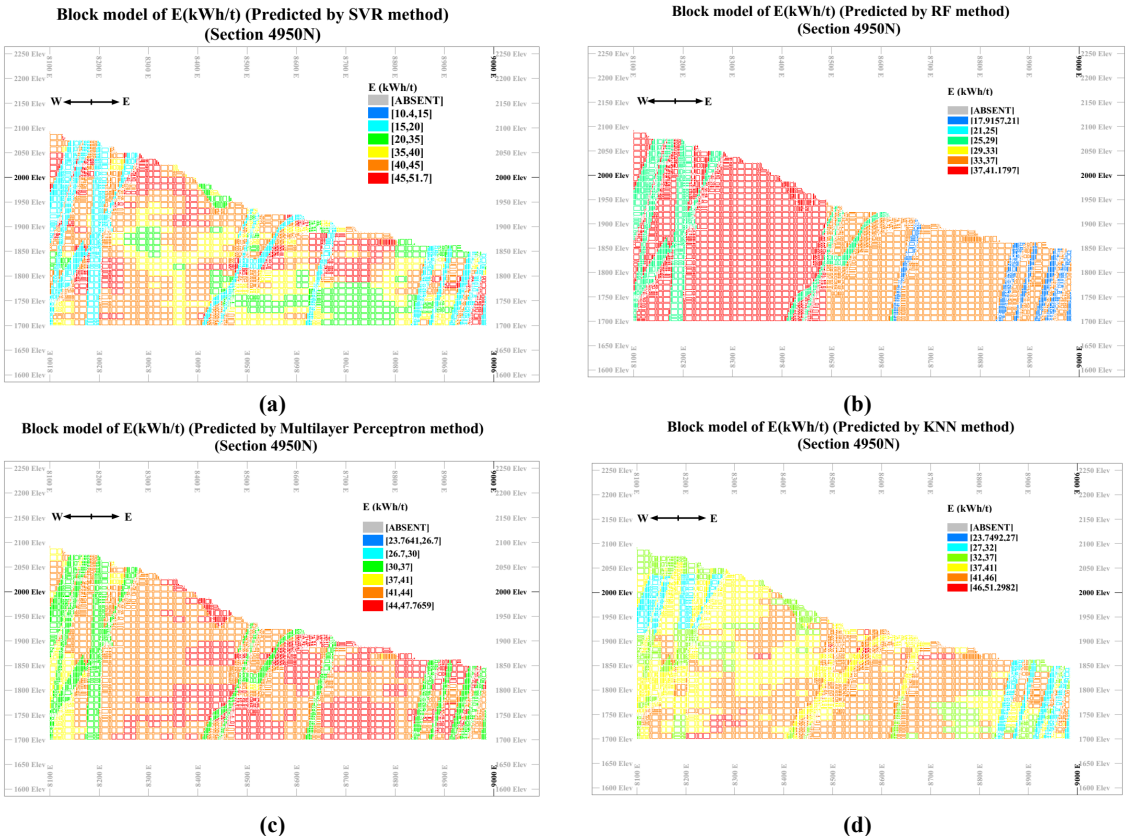


Figure 6. Cross-section of block model for SPI test created by: a) SVR, b) RF, c) Multilayer Perceptron and d) KNN methods.

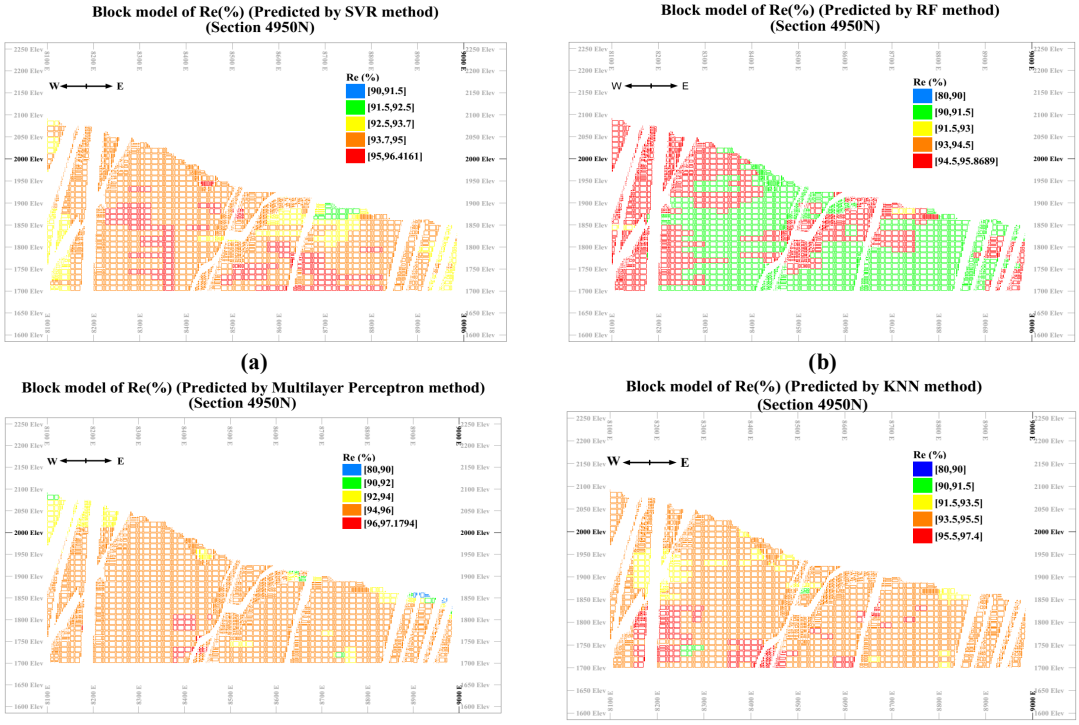


Figure 7. Cross-section of block model for Cu recovery, created by: a) SVR, b) RF, c) Multi-layer perceptron and d) KNN methods.

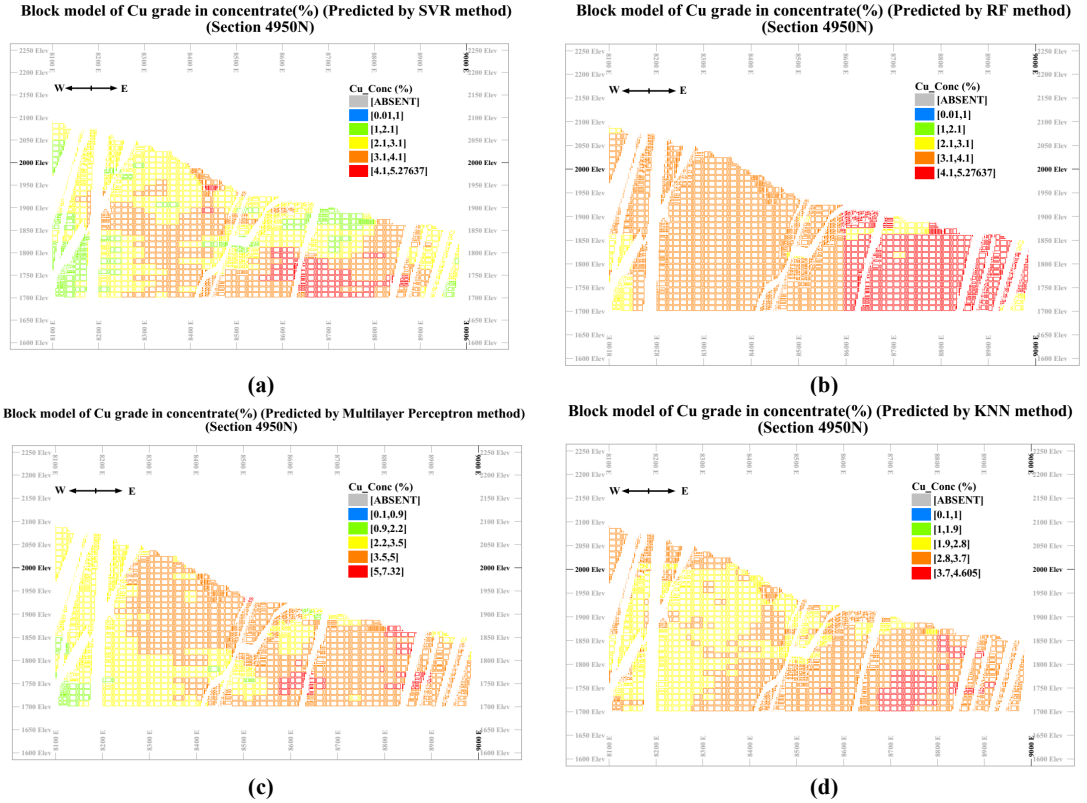


Figure 8. Cross-section of block model for Cu grade in concentrate created by: a) SVR, b) RF, c) Multilayer Perceptron and d) KNN methods.

In the context of RSD, ML methods are deemed satisfactory when the upper bound limit of RSD is below 25% [51]. On the other hand, RSD values below 5% indicate the excellence of the machine-learning method for spatial modeling of the assessed process property [52]. Based on the information presented in Table 5 and Figure 9-a, the RSD offers a distinct advantage when comparing the results obtained for the modelled process properties. Across all methods, the predictions for Cu recoveries outperform those for Cu grade conc. and SPI. Excellent predictions were achieved for Cu recoveries, as all the methods demonstrated a RSD below 5%. This indicates a high level of accuracy and consistency in the predicted Cu recoveries across the different methods employed. Based on the findings, it can be inferred that any of the machine-learning methods tested can be utilized for spatial modeling of Cu recovery. However, when comparing the RMSE, MAE, and R^2, it becomes evident that the most suitable method for this particular task is SVR. The lower values of RMSE and MAE, along with the higher R^2, indicate that SVR outperforms the other tested methods in terms of accuracy and predictive performance for modeling Cu recovery. Therefore, SVR is recommended as the most appropriate choice for spatial modeling of Cu recovery based on these evaluation metrics.

Contrary to the results observed for Cu recovery, the RSD for the SVR model in the SPI test is the highest among all the methods. On the other hand, the KNN method exhibits lower values of RMSE, MAE, and RSD, while demonstrating higher values of the coefficient of determination. These findings suggest that the KNN method is more appropriate for the SPI test in terms of its superior accuracy and predictive performance, as indicated by the lower error metrics and higher R^2 value. Therefore, for modeling SPI, the KNN method is recommended over SVR based on these evaluation measures.

Similar to the modeling of Cu recovery, the SVR method proves to be effective in modeling Cu grade in concentrate. While the MLP method also exhibits high values of the R^2, it has higher levels of RSD compared to SVR. This suggests that the SVR method is more reliable and consistent in capturing the variations and accurately predicting Cu grade in concentrate. Despite the high R^2 values of the MLP method, the higher RSD implies a greater degree of uncertainty and variability in the predictions. Therefore, when considering both accuracy and reliability, the SVR method is considered more suitable for modeling Cu grade in concentrate.

Table 5. Comparison of evaluation metrics for MLAs in GM.

MLAs	Response variable	R^2	MAE	RMSE	RSD
KNN	***SPI***	0.46	7.37	9.16	13.81
	Re	0.47	2.04	2.56	2.12
	Cu Concentrate	0.52	0.98	1.24	15.20
MLP	***SPI***	0.28	10.23	12.02	15.43
	Re	0.76	1.44	1.72	2.28
	Cu Concentrate	0.87	0.39	0.52	26.60
RF	***SPI***	0.41	7.45	10.19	26.26
	Re	0.09	2.27	3.13	2.34
	Cu Concentrate	0.56	1.02	1.19	30.02
SVR	***SPI***	0.55	8.22	9.11	44.27
	Re	0.66	1.05	1.23	0.80
	Cu Concentrate	0.61	0.60	0.79	18.58

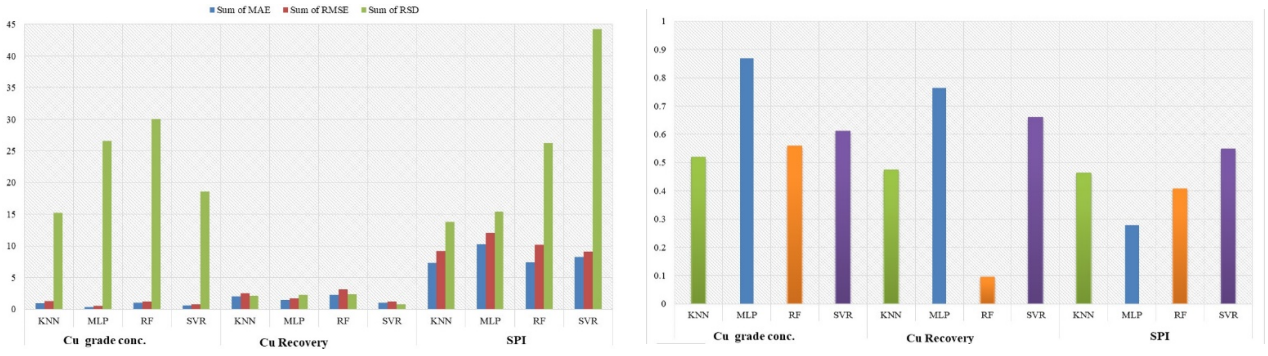


Figure 9. Comparison of geometallurgical models performance for different MLAs: a) RMSE, MAE and RSD, b) R^2.

6. Conclusions

Indeed, GM represents the final stage of a geometallurgical program. This modeling process involves the integration and analysis of geological, grade, and metallurgical data to develop a comprehensive understanding of the deposit's behavior. By employing statistical analysis, and MLA modeling techniques, the spatial distribution and variability of crucial variables can be predicted. The ultimate goal of this final step is to

provide valuable insights for decision-making in mining operations, resource estimation, and process optimization, thereby maximizing the efficiency and profitability of the overall geometallurgical program. When conducting GM for Cu porphyry deposits, it is essential to take into account a range of important variables and parameters. These include ore composition, mineralogy, physical properties, and flotation recovery, all of which exert a substantial influence on the behavior of the ore during mineral processing operations. By comprehensively considering these factors, a more accurate and effective understanding of the ore's characteristics and response to processing can be achieved.

In this paper, the main objective was to compare MLAs to find an efficient way of constructing spatial models for metallurgical response. At the Sungun Copper mine in Iran, a comprehensive assessment of the metallurgical performance was conducted by measuring several process properties. These properties included SPI (SAG Mill Power Index), Copper recovery, and Copper grade in concentrate. To predict these properties, sample coordinates, grades, lithology, mineralization zone, and alteration were utilized as primary data inputs. By incorporating these diverse variables, a holistic understanding of the metallurgical behavior and performance at Sungun Copper mine was achieved, facilitating informed decision-making and optimization of the mining and processing operations. When employing MLA for spatial modeling, it is imperative to evaluate not only the error metrics but also the RSD to ensure accurate and reliable results. In this case, the predictions for Cu recoveries stand out as they consistently exhibit an RSD below 5% across all methods, indicating excellent accuracy and reliability. Contrary to the results observed for Cu recovery, the RSD for the SVR model in the SPI test is the highest among all the methods. On the other hand, the KNN method exhibits lower values of RMSE, MAE, and RSD, while demonstrating higher values of the R^2. These findings suggest that the KNN method is more appropriate for the SPI test in terms of its superior accuracy and predictive performance, as indicated by the lower error metrics and higher R^2 value.

Acknowledgments

This research forms an integral part of a Ph.D. project, which has received support from the National Iranian Copper Industries Co. (NICICO). The authors would like to express their gratitude to all the personnel of Sungun mine for their valuable assistance in providing the necessary data and conducting the sampling process. Their contributions have been instrumental in the successful execution of this study.

References

[1]. Deutsch, C. (2013). *Geostatistical modelling of geometallurgical variables—Problems and solutions.* Paper presented at the Proceedings of the International Geometallurgy Conference, Brisbane, Australia.

[2]. Gordon, H. J. J. (2019). *A mineralogical approach to quantifying ore variability within a polymetallic Cu-Pb-Zn Broken Hill-type deposit and its implications for geometallurgy.* Stellenbosch: Stellenbosch University.

[3]. Lishchuk, V. (2016). Geometallurgical programs–critical evaluation of applied methods and techniques.

[4]. Sepúlveda Escobedo, E. M. (2018). *Quantification of uncertainty of geometallurgical variables for mine planning optimisation.*

[5]. Vann, J., Jackson, J., Coward, S., & Dunham, S. (2011). The geomet curve–a model for implementation of geometallurgy. *First AusIMM International Geometallurgy*, 1-10.

[6]. Walters, S., & Kojovic, T. (2006). Geometallurgical mapping and mine modelling (GEMIII)-the way of the future.

[7]. Williams, S. R., & Richardson, J. (2004). *Geometallurgical Mapping: A new approach that reduces technical risk.* Paper presented at the Proceedings 36th Annual Meeting of the Canadian Mineral Processors.

[8]. Lishchuk, V., Lamberg, P., & Lund, C. (2015). *Classification of geometallurgical programs based on approach and purpose.* Paper presented at the SGA 2015: 24/08/2015-27/08/2015.

[9]. Dominy, S., Murphy, B., & Gray, A. (2011). *Characterisation of gravity amenable gold ores—Sample representivity and determination methods.* Paper presented at the Proceedings International Geometallurgy Conference; The Australasian Institute of Mining and Metallurgy: Melbourne, Australia.

[10]. Lishchuk, V., & Pettersson, M. (2021). The mechanisms of decision-making when applying geometallurgical approach to the mining industry. *Mineral Economics, 34*, 71-80.

[11]. Dominy, S. C., O'Connor, L., Parbhakar-Fox, A., Glass, H. J., & Purevgerel, S. (2018). Geometallurgy—A route to more resilient mine operations. *Minerals, 8*(12), 560.

[12]. Garrido, M., Sepúlveda, E., Ortiz, J. M., Navarro, F., & Townley, B. (2018). *A methodology for the simulation of synthetic geometallurgical block models of porphyry ore bodies.* Paper presented at the Procemin 14th International Mineral Processing Conference (PROCEMIN). 5th International Seminar on Geometallurgy (GEOMET). Santiago, Chile.

[13]. Hunt, J., Berry, R., Bradshaw, D., Triffett, B., & Walters, S. (2014). Development of recovery domains: Examples from the Prominent Hill IOCG deposit, Australia. *Minerals Engineering, 64*, 7-14.

[14]. Hunt, J., Kojovic, T., & Berry, R. (2013). Estimating comminution indices from ore mineralogy, chemistry and drill core logging.

[15]. Lamberg, P. (2011). *Particles-the bridge between geology and metallurgy.* Paper presented at the Konferens i mineralteknik 2011: 08/02/2011-09/02/2011.

[16]. Sepulveda, E., Dowd, P., Xu, C., & Addo, E. (2017). Multivariate modelling of geometallurgical variables by projection pursuit. *Mathematical Geosciences, 49*, 121-143.

[17]. Patnaik, S., Yang, X.-S., & Sethi, I. K. (2019). Advances in Machine Learning and Computational Intelligence Proceedings of ICMLCI 2019. *Proceedings of ICMLCI*, 1.

[18]. Suthaharan, S. (2016). Machine learning models and algorithms for big data classification. *Integr. Ser. Inf. Syst, 36*, 1-12.

[19]. Oliver, S., & Willingham, D. (2016). *Maximise orebody value through the automation of resource model development using machine learning.* Paper presented at the Proceedings of the Third AusIMM International Geometallurgy Conference, Perth, Australia.

[20]. Sagar, D., Cheng, Q., & Agterberg, F. (2018). *Handbook of mathematical geosciences: fifty years of IAMG*: Springer Nature.

[21]. Afzal, P., Farhadi, S., Boveiri Konari, M., Shamseddin Meigooni, M., & Daneshvar Saein, L. (2022). Geochemical anomaly detection in the Irankuh District using Hybrid Machine learning technique and fractal modeling. *Geopersia, 12*(1), 191-199.

[22]. Farhadi, S., Afzal, P., Boveiri Konari, M., Daneshvar 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.

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

[24]. Abbaszadeh, M., Hezarkhani, A., & Soltani-Mohammadi, S. (2013). An SVM-based machine learning method for the separation of alteration zones in Sungun porphyry copper deposit. *Geochemistry, 73*(4), 545-554.

[25]. Ebdali, M., & Hezarkhani, A. (2024). A comparative study of decision tree and support vector machine methods for gold prospectivity mapping. *Mineralia Slovaca, 56*(2).

[26]. Saremi, M., Maghsoudi, A., Hoseinzade, Z., & Mokhtari, A. R. (2024). Data-driven AHP: A novel method for porphyry copper prospectivity mapping in the Varzaghan District, NW Iran. *Earth Science Informatics, 17*(6), 5063-5078.

[27]. Saljoughi, B. S., & Hezarkhani, A. (2024). A Comparative Analysis of Artificial Neural Network (ANN) and Gene Expression Programming (GEP) Data-driven Models for Prospecting Porphyry Cu Mineralization; Case Study of Shahr-e-Babak Area, Kerman Province, SE Iran. *Journal of Mining and Environment, 15*(2), 761-790.

[28]. Rajabinasab, B., & Asghari, O. (2019). Geometallurgical domaining by cluster analysis: Iron ore deposit case study. *Natural Resources Research, 28*, 665-684.

[29]. Lishchuk, V., Lund, C., & Ghorbani, Y. (2019). Evaluation and comparison of different machine-learning methods to integrate sparse process data into a spatial model in geometallurgy. *Minerals Engineering, 134*, 156-165.

[30]. Coward, S., Vann, J., Dunham, S., & Stewart, M. (2009). *The primary-response framework for geometallurgical variables.* Paper presented at the Seventh international mining geology conference.

[31]. Carrasco, P., Chilès, J.-P., & Séguret, S. A. (2008). *Additivity, metallurgical recovery, and grade.* Paper presented at the 8th international Geostatistics Congress.

[32]. Keeney, L., & Walters, S. (2011). *A methodology for geometallurgical mapping and orebody modelling.* Paper presented at the GeoMet 2011-1st AusIMM International Geometallurgy Conference 2011.

[33]. Ameh.mxen, P. (2003). *The application of the SAG POWER INDEX to ore body hardness characterization for the design and optimization of autogenous grinding circuits.* (Master). McGiII University, Montreal.

[34]. Starkey, J., & Dobby, G. (1996). Application of the Minnovex SAG power index at five Canadian SAG plants. *Proceeding Autogenous and Semi-Autogenous Grinding*, 345-360.

[35]. Bahrami, A., Ghorbani, Y., Sharif, J. A., Kazemi, F., Abdollahi, M., Salahshur, A., & Danesh, A. (2021). A geometallurgical study of flotation

performance in supergene and hypogene zones of Sungun copper deposit. *Mineral Processing and Extractive Metallurgy, 130*(2), 126-135.

[36]. Amankwaa-Kyeremeh, B., McCamley, C., Zanin, M., Greet, C., Ehrig, K., & Asamoah, R. K. (2023). Prediction and Optimisation of Copper Recovery in the Rougher Flotation Circuit. *Minerals, 14*(1), 36.

[37]. Boisvert, J. B., Rossi, M. E., Ehrig, K., & Deutsch, C. V. (2013). Geometallurgical modeling at Olympic dam mine, South Australia. *Mathematical Geosciences, 45*, 901-925.

[38]. Macmillan, E., Ehrig, K., Liebezeit, V., Kittler, P., Lower, C. (2011). *Use of geometallurgy to predict tailings leach acid consumption at Olympic Dam*. Paper presented at the The First AUSIMM International Geometallurgy Conference, Brisbane.

[39]. Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4): Springer.

[40]. Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2): Springer.

[41]. Meng, M., & Zhao, C. (2015). Application of support vector machines to a small-sample prediction. *Advances in Petroleum Exploration and Development, 10*(2), 72-75.

[42]. Breiman, L. (2001). Random forests. *Machine learning, 45*, 5-32.

[43]. Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*: MIT press.

[44]. Hezarkhani, A., Williams-Jones, A., & Gammons, C. (1999). Factors controlling copper solubility and chalcopyrite deposition in the Sungun porphyry copper deposit, Iran. *Mineralium deposita, 34*, 770-783.

[45]. Hezarkhani, A. (1997). Physicochemical controls on alteration and copper mineralization in the Sungun porphyry copper deposit, Iran.

[46]. Calagari, A. A. (2003). Stable isotope (S, O, H and C) studies of the phyllic and potassic–phyllic alteration zones of the porphyry copper deposit at Sungun, East Azarbaidjan, Iran. *Journal of Asian Earth Sciences, 21*(7), 767-780.

[47]. Asghari, O., & Hezarkhani, A. (2008). Appling discriminant analysis to separate the alteration zones within the Sungun porphyry copper deposit. *Journal of Applied Sciences, 24*, 4472-4486.

[48]. Hezarkhani, A., & Williams-Jones, A. E. (1998). Controls of alteration and mineralization in the Sungun porphyry copper deposit, Iran; evidence from fluid inclusions and stable isotopes. *Economic Geology, 93*(5), 651-670.

[49]. Nikfarjam, M., Hezarkhani, A., & Aziz-Afshari, F. (2022). Geological domaining at Sungun porphyry copper deposit using cluster analysis. *Global Journal of Computer Sciences: Theory and Research, 12*, 78-92.

[50]. Mery, N., et al., *Geostatistical modeling of the geological uncertainty in an iron ore deposit.* Ore Geology Reviews, 2017. **88**: p. 336-351.

[51]. David, D. (2013). *Geometallurgical guidelines for miners, geologists and process engineers– discovery to design.* Paper presented at the The Second AusIMM International Geometallurgy Conference, Brisbane, QLD, Australia.

[52]. Grubbs, F. E. (1969). Procedures for Detecting Outlying Observations in Samples. *Technometrics, 11*(1), 1-21. doi:10.1080/00401706.1969.10490657.

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

تحلیل مقایسه‌ای روش‌های رگرسیون یادگیری ماشین برای مدل‌سازی ژئومتالورژیکی در کانسار مس پورفیری سونگون

میثم نیک فرجام[۱]، اردشیر هزارخانی[۱*]، فرهاد عزیزافشاری[۲] و حمیدرضا گلچین[۲]

۱. بخش مهندسی معدن، دانشگاه صنعتی امیرکبیر، تهران، ایران
۲. شرکت ملی مس ایران، معدن مس سونگون، آذربایجان شرقی، ایران

اطلاعات مقاله

تاریخ ارسال: ۲۰۲۴/۰۹/۱۸
تاریخ داوری: ۲۰۲۵/۰۳/۲۷
تاریخ پذیرش: ۲۰۲۵/۰۵/۱۶

DOI: 10.22044/jme.2025.15088.2885

کلمات کلیدی

مدل‌سازی ژئومتالورژی
خواص فرآیند
شاخص توان نیمه خودزا (SPI)
الگوریتم‌های یادگیری ماشین
معدن مس سونگون

چکیده

مدل‌سازی ژئومتالورژیکی نقش مهمی در بخش معدن دارد و امکان درک جامع از روابط پیچیده بین عوامل زمین‌شناسی و متالورژیکی را فراهم می‌کند. این مطالعه بر ارزیابی متغیرهای متالورژیکی در معدن مس سونگون در ایران تمرکز دارد و با اندازه‌گیری و پیش‌بینی ویژگی‌های فرآوری، از جمله شاخص آسیای نیمه‌خودشکن (SPI)، بازیابی (Re) و عیار، انجام می‌شود. برای غلبه بر محدودیت‌های ناشی از ویژگی جمع‌پذیری روش‌های زمین آماری، با هدف بهبود تصمیم‌گیری و بهینه‌سازی عملیات معدن، از الگوریتم‌های یادگیری ماشین برای بهبود مدل‌سازی پیش‌بینی استفاده شد. در این پژوهش ورودی‌ها را داده‌هایی نظیر مختصات نمونه‌ها، عیار، لیتولوژی، زون‌های کانی‌سازی و دگرسانی تشکیل داده است تا دقت و قابلیت اطمینان روش‌های مختلف رگرسیون یادگیری ماشین را ارزیابی کند. انحراف معیار نسبی (RSD) به عنوان یک معیار مهم برای مقایسه دقت ویژگی‌های پیش‌بینی شده فرآوری مورد استفاده قرا گرفته است. افزون بر این، معیارهای ارزیابی مانند خطای جذر میانگین مربعات (RMSE)، میانگین خطای مطلق (MAE) و ضریب تعیین (R2) برتری روش‌های مدل‌سازی مورد استفاده را در سناریوهای مختلف بررسی می‌کنند. روش Kنزدیک‌ترین همسایه (KNN) دقت بالاتر، معیارهای خطای کمتر و R2 بالاتری را برای مدل‌سازی آزمون SPI نشان می‌دهد. برای مدل‌سازی عیار مس در کنسانتره، رگرسیون بردار پشتیبان (SVR) مؤثر و قابل اعتماد است و از روش پرسپترون چندلایه (MLP) پیشی می‌گیرد. با وجود R2 بالای روش MLP، RSD بالاتر آن نشان‌دهنده عدم قطعیت و تغییرپذیری بیشتر در پیش‌بینی‌ها است. بنابراین، SVR برای مدل‌سازی عیار مس در کنسانتره مناسب‌تر در نظر گرفته می‌شود.

نویسنده مسئول مکاتبات: ardehez@aut.ac.ir (ا. هزارخانی)