






Article

Metallurgical Copper Recovery Prediction Using Conditional Quantile Regression Based on a Copula Model

Heber Hernández ^{1,*}, Martín Alberto Díaz-Viera ², Elisabete Alberdi ³, Aitor Oyarbide-Zubillaga ⁴
and Aitor Goti ⁴

¹ Facultad de Ingeniería y Arquitectura, Universidad Central de Chile, Santiago 8370178, Chile

² Instituto Mexicano del Petróleo, Eje Central Lázaro Cárdenas No. 152, Ciudad de México 07730, Mexico; mdiazv@imp.mx

³ Department of Applied Mathematics, University of the Basque Country UPV/EHU, 48013 Bilbao, Spain; elisabete.alberdi@ehu.es

⁴ Department of Mechanics, Design and Organization, University of Deusto, 48007 Bilbao, Spain; aitor.oyarbide@deusto.es (A.O.-Z.); aitor.goti@deusto.es (A.G.)

* Correspondence: heber.hernandez@ucentral.cl

Abstract: This article proposes a novel methodology for estimating metallurgical copper recovery, a critical feature in mining project evaluations. The complexity of modeling this nonadditive variable using geostatistical methods due to low sampling density, strong heterotopic relationships with other measurements, and nonlinearity is highlighted. As an alternative, a copula-based conditional quantile regression method is proposed, which does not rely on linearity or additivity assumptions and can fit any statistical distribution. The proposed methodology was evaluated using geochemical log data and metallurgical testing from a simulated block model of a porphyry copper deposit. A highly heterotopic sample was prepared for copper recovery, sampled at 10% with respect to other variables. A copula-based nonparametric dependence model was constructed from the sample data using a kernel smoothing method, followed by the application of a conditional quantile regression for the estimation of copper recovery with chalcocite content as secondary variable, which turned out to be the most related. The accuracy of the method was evaluated using the remaining 90% of the data not included in the model. The new methodology was compared to cokriging placed under the same conditions, using performance metrics RMSE, MAE, MAPE, and R^2 . The results show that the proposed methodology reproduces the spatial variability of the secondary variable without the need for a variogram model and improves all evaluation metrics compared to the geostatistical method.

Keywords: metallurgical copper recovery; copula model; conditional quantile regression; kernel smoothing; collocated cokriging



Citation: Hernández, H.; Díaz-Viera, M.A.; Alberdi, E.; Oyarbide-Zubillaga, A.; Goti, A. Metallurgical Copper Recovery Prediction Using Conditional Quantile Regression Based on a Copula Model. *Minerals* **2024**, *14*, 691. <https://doi.org/10.3390/min14070691>

Academic Editors: Krzysztof Skrzypkowski, René Gómez, Fhatuwani Sengani, Derek B. Apel, Faham Tahmasebinia and Jianhang Chen

Received: 20 May 2024

Revised: 24 June 2024

Accepted: 25 June 2024

Published: 1 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Metallurgical recovery, in the context of mineral mining, refers to the percentage of valuable metal extracted from the ore during the processing or beneficiation stage [1,2]. It is a crucial metric in the mining industry as it indicates the efficiency of the extraction process and ultimately affects the profitability of the mining operation.

From a mineral processing perspective, contemporary treatments for copper include the flotation of sulfide ores, leaching for oxide ores, and a hybrid approach that integrates flotation and magnetic separation for certain mixed ores [3]. Madenova and Madani (2021) [4] define metallurgical recovery as a vital geometallurgical variable for mine planning, representing a response to the processing plant design and the geological characteristics of the ore.

The process of extracting valuable metals from ore typically involves several stages, including crushing, grinding, concentration, and refining. Metallurgical recovery measures

the effectiveness of these processes in separating and concentrating the valuable metal from the ore.

The calculation of metallurgical recovery involves comparing the amount of metal recovered from the ore to the total amount of metal present in the ore. This is often expressed as a percentage, where a higher percentage indicates a more efficient extraction process.

Factors influencing metallurgical recovery include the mineralogy and composition of the ore, the efficiency of the processing equipment and techniques used, and the expertise of the personnel involved in the operation. Improving metallurgical recovery is a key focus area for mining companies seeking to optimize their operations and maximize the value of their mineral resources.

To maximize profit, it is crucial to have a reliable estimation model for all variables involved, both primary (geological properties, mineral grades, densities, contaminants, etc.) and response variables (metallurgical recovery, bond work index, grindability index, processing capacity, milling performance, etc.) [5]. The combination of a geological model and metallurgical data is currently known as a geometallurgical model [6].

Metallurgical recovery is a critical feature in the evaluation and exploitation of mining projects, as it directly influences net economic benefit. This variable is expressed as a percentage and represents the yield of mineral processing, in the case of copper sulfides, by flotation along the mining value chain [7].

Incorporating metallurgical recovery into mine planning poses a challenge for resource modelers and mine planners. In most projects, the lack of proper collection and analysis of geometallurgical data leads to unreliable metallurgical response models [8]. Samples with this information are often scarce, costly, and highly heterotopic compared to primary variables [9,10]. However, these primary variables are known to be useful indicators for predicting metallurgical responses [11,12].

In the context of copper metallurgical recovery, a heterotopic sample refers to a sample taken from a location within a mineral deposit that is geologically distinct from the primary ore body being targeted for extraction.

For example, in a copper mining operation, the primary ore body may consist of a specific geological formation or vein where copper minerals are concentrated. A heterotopic sample, in this case, could be taken from a nearby area where copper mineralization occurs but in a different geological setting. This could include samples from adjacent rock formations, secondary veins, or areas with different mineralogical characteristics.

Analyzing heterotopic samples is important in copper metallurgical recovery because they can provide insights into the variability of copper mineralization within a mining area. Understanding the distribution and characteristics of copper mineralization in heterotopic samples can help optimize mining processes, improve recovery rates, and inform mine planning and development strategies.

Modeling metallurgical recovery is usually carried out using geostatistical techniques and, more recently, by machine learning methods. Machine learning methods seek to predict geometallurgical response variables using assay and mineralogy data [13–17]; however, they require a large number of variables and a considerable amount of data to consolidate a robust model [18]. On the other hand, geostatistics requires initially defining domains based on geometallurgical attributes [19] and then estimating indirectly, that is, by seeking mathematical arrangements forced to meet the required assumptions. Techniques like kriging are not usually recommended, as they can generate biased results due to the non-additive nature of metallurgical recovery [20]. It has been observed that the weighted average of two sample values is not a good estimator of the corresponding value in the blend [21,22]. Additionally, cokriging and its variants have difficulties being applicable, mainly due to the subjectivity in modeling variograms with little information and nonlinear or complex dependency relationships with mineral grades and other geochemical variables [23].

Given the problems and limitations mentioned earlier, the application of copula-based methods emerges as a promising alternative in the mining industry [24]. As is usual in statistics and geostatistics, two approaches have been developed: one for estimation and

the other for simulation. In particular, several geostatistical methods based on copulas have been published that have been successful in Earth sciences applications [25–29]. But above all, there are multiple developments in the sphere of finance using an estimation approach known as quantile regression based on copulas [30–32].

A copula is a multivariate cumulative distribution function for which the marginal probability distribution of each variable is uniform on the interval [0, 1]. They are functions that describe the underlying dependence between random variables. Sklar's theorem [33] states that any multivariate joint distribution can be expressed in terms of univariate marginal distribution functions and a copula that describes the dependency structure between the variables [34,35].

This research evaluates the performance of the conditional quantile regression method (CQRM) compared to the classical geostatistical collocated cokriging method (CCM) for modeling copper metallurgical recovery based on a predominantly measured geological attribute.

The structure of the paper is as follows. In the Section 2, the problem statement is established. The Section 3 presents the collocated cokriging and conditional quantile regression methods, and the general methodology of their application. The Section 4 gives a description of the dataset used in the comparative study. The Section 5 shows the results of the application of both methods to the case study. In the Section 6, the comparison of the performance of the two methods is discussed, and, finally, in the Section 7, the conclusions and future work are given.

2. Problem Statement

In a basic mining exploitation unit to calculate mining profit, Equation (1), metallurgical recovery, ore grade, and metal price are critical variables, all subject to significant uncertainty, as can be inferred from its own definition [36,37]:

$$Profit = R_{rc}^{Cu} \cdot P_c^{Cu} \cdot G^{Cu} \cdot T^{Cu} - C_c^{Cu} \quad (1)$$

where P_c^{Cu} is the price of the copper concentrate, G^{Cu} is the copper grade, T^{Cu} is the mineral tonnage, and C_c^{Cu} are the costs associated with the production, processing, and sale of the copper concentrate.

Due to the low sampling density, their strong heterotopic relationship with other mineral deposit measurements, and the nonlinearity in these relationships, geostatistical modeling appears to be a complex alternative.

Formally, metallurgical recovery can be defined for a copper mine by the following expression [8]:

$$R_{rc}^{Cu} = \frac{m_c^{Cu}}{m^{Cu}} = x_c^{Cu} f_c \frac{1}{x^{cu}} \quad (2)$$

where R_{rc}^{Cu} is the copper recovery, m_c^{Cu} is the mass of copper in the concentrate (i.e., the amount of recovered copper), and m^{Cu} is the initial mass of copper in the feed. x_c^{Cu} is the copper grade in the concentrate, x^{cu} is the feed grade, and f_c is the fraction of mass recovered.

Small variations in metallurgical recovery estimation will impact the profit valuation, which defines the material's destinations in strategic planning, whether as processed ore in the plant, waste rock extracted for deposition in dumps, or low-grade ore for stockpiling. Given the importance of this variable for downstream processes such as block model optimization, mine design, and mine life planning, it is essential to seek the best practices for its prediction.

3. Methodology

In this paper, a comparison of the conditional quantile regression method (CQRM) is made with respect to the traditional collocated cokriging method (CCM) in terms of accuracy and performance. To apply both methods, there are a series of steps that are common, as is shown in the general methodological workflow of Figure 1.

The first step, consisting of the exploratory analysis of the data, is standard for any statistical procedure and consists of a summary of its statistics and the probability distribution graphics with its boxplots and histograms. Also, an evaluation of the impact of the presence of outliers in the data sample on the statistics is also carried out.

In the variable selection step in both cases, the variable pair that has the greatest dependence is sought, but the difference is that for CCM it must be a linear dependence that should be estimated with the Pearson correlation coefficient, while for CQRM it is recommendable to use a more robust measure of dependence by using Spearman or Kendall rank correlation coefficients.

In the modeling part, different models are built. For CCM it is a spatial correlation model with the variogram of the primary variable, while for CQRM it is the dependence model that consists of the estimation and fitting of the marginal distributions of each variable as well as the copula.

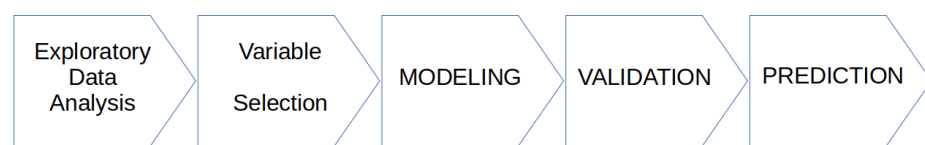


Figure 1. General methodological workflow.

The validation of the models is carried out for CCM with the usual cross-validation method, i.e., a leave one out method [38], while CQRM is validated performing an estimation with the same data used to build it. The quality of each method is evaluated in terms of performance metrics, such as root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and determination coefficient (R^2).

Finally, the prediction of copper metallurgical recovery conditioned by a secondary attribute is performed at each point where there are no sample values.

3.1. Collocated Cokriging Method

Collocated cokriging is a geostatistical method used for spatial interpolation or estimation of a target variable at unsampled locations based on available data from sampled locations [38]. This method is particularly useful when dealing with multiple correlated variables or when auxiliary information is available. In collocated cokriging, the relationship between the target variable and auxiliary variables is modeled using the concept of spatial dependence, which assumes that nearby locations tend to have similar values. The method estimates the target variable at an unsampled location by combining information from both the target variable and auxiliary variables at nearby sampled locations (see Appendix B).

The usual steps involved in the application of collocated cokriging method are the following:

- **Data collection:** Collect data on the target variable and auxiliary variables from sampled locations within the study area.
- **Spatial correlation analysis:** Assess the spatial correlation or dependence between the target variable and auxiliary variables using variograms or covariance functions.
- **Modeling:** Model the spatial dependence structure between the target variable and auxiliary variables using geostatistical techniques such as ordinary kriging.
- **Validation:** Validate the accuracy of the predictions using cross-validation or comparison with independent data, if available.
- **Prediction:** Estimate the value of the target variable at unsampled locations by combining information from nearby sampled locations and auxiliary variables, taking into account their spatial correlation.

Collocated cokriging offers advantages over traditional kriging methods by incorporating additional information from auxiliary variables, which can improve the accuracy of predictions, especially in areas with limited or sparse data coverage for the target variable.

3.2. Conditional Quantile Regression Method

Conditional quantile regression based on copulas is a statistical method used for modeling the relationship between variables, particularly when dealing with non-normal or skewed distributions, and when there are complex dependencies among variables. Copulas are mathematical functions that describe the dependence structure between random variables, independent of their marginal distributions. They capture the joint distribution of variables without making assumptions about their individual distributions. Copulas allow the modeling of both linear and nonlinear dependencies, making them useful for capturing complex relationships between variables.

Conditional quantile regression extends the concept of linear regression by estimating conditional quantiles of the response variable given the values of predictor variables. Unlike ordinary least squares regression, which models the conditional mean of the response variable, quantile regression models different quantiles (e.g., median, upper/lower quantiles), providing a more comprehensive understanding of the conditional distribution of the response variable. In quantile regression based on copulas, copulas are used to model the dependence structure between variables, while quantile regression is applied to estimate the conditional quantiles of the response variable given the values of predictor variables. This approach allows the capturing of the joint distribution and conditional distributions of variables simultaneously, taking into account complex dependencies among them.

Conditional quantile regression based on copulas offers several advantages:

- Flexibility: It can model nonlinear dependencies and account for heteroscedasticity.
- Robustness: It is robust to outliers and non-normality in the data.
- Interpretability: It provides insights into how different quantiles of the response variable are affected by changes in the predictor variables.
- Tail behavior: It can capture extreme events or tail behavior in the conditional distributions of variables.

This method is widely used in finance, economics, environmental science, and other fields where understanding the relationship between variables across different quantiles is essential. It can be used for risk management, forecasting, and decision making under uncertainty. Overall, quantile regression based on copulas is a powerful tool for modeling complex dependencies and understanding the conditional distribution of variables, especially in situations where traditional regression methods may not be adequate.

In this article, a nonparametric copula by the kernel smoothing method [39] is used, given its good adaptability to any type of distribution and computational efficiency. In particular, kernel smoothing, a type of weighted moving average method, is applied for probability density estimation, that is, to estimate the probability density function of a random variable based on kernels as weights, where the term kernel in this context means a window function.

A more detailed explanation of copula theory and the quantile regression method can be found in Appendix A. In particular, additional explanations about the kernel smoothing method can be found in Appendix A.2.

4. Dataset Description

With the purpose of carrying out the comparison between the collocated cokriging and conditional quantile regression methods with mining data, a dataset extracted from a synthetic porphyry copper deposit was selected, which was published in [19], and is openly available for academic use. From the three-dimensional block model, the level 2080 is chosen to configure a 2D space and we use it as experimental data for this application. This dataset consists of 3479 (71×49) cells of 20×20 meters, each spatially georeferenced by their X (easting) and Y (northing) coordinates. Each cell contains geochemical records

of 1—clays, 2—chalcocite, 3—bornite, 4—chalcopyrite, 5—tennantite, 6—molybdenite, 7—pyrite, 8—copper (Cu), 9—molybdenum (Mo), and 10—arsenic (As), as well as the results of metallurgical tests for 11—copper recovery and 12—bond work index.

This exercise consists of extracting, in a random and spatially uniform manner, a sample equivalent to 10% of the total available information of the metallurgical recovery, constructing a strongly heterotopic case with respect to the other variables. Figure 2 shows the copper recovery maps for the complete dataset on the left side and the corresponding map for the 10% sampling on the right side, respectively.

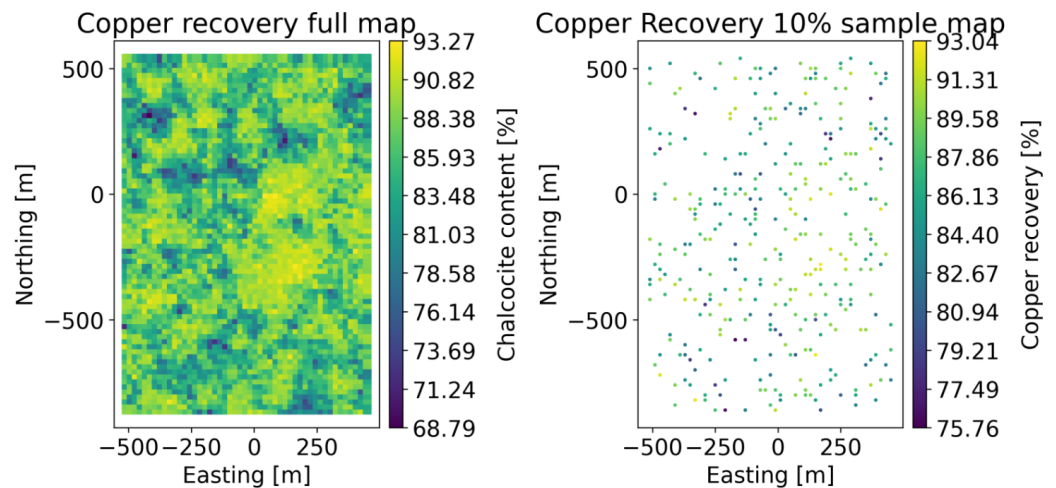


Figure 2. Copper recovery for the 100% (in the left) and 10% (in the right) sample maps, respectively.

The comparison of the CCM and CQRM methods is carried out by applying them to the rest of the data corresponding to 90% of the dataset, allowing a direct comparison between the real information and the results obtained by the two methods under the same heterotopic sampling conditions and validation metrics.

A comparative statistical summary between the full (100%) dataset and the 10% sample for copper recovery is shown in Table 1 and Figures 3 and 4, respectively. It can be seen that the 10% subset of data is statistically equivalent to the total dataset in terms of the statistical values and probability distribution. Hereinafter, the subset corresponding to 10% of the dataset will be referred to as “10% sample”.

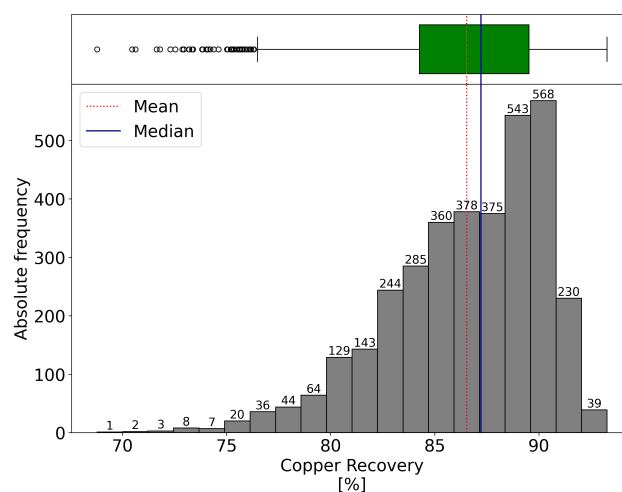


Figure 3. Copper recovery histogram and boxplot for the full (100%) dataset.

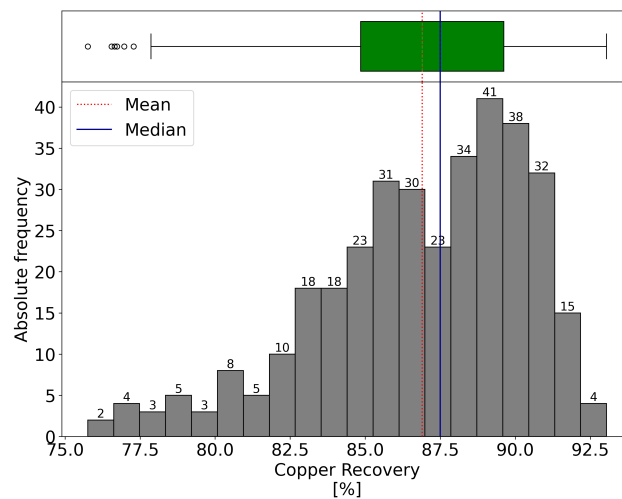


Figure 4. Copper recovery histogram and boxplot for 10% sample dataset.

Table 1. Copper recovery statistics summary for the total and 10% sample, respectively.

Statistics	Total Sample	10% Sample
Size	3479	347
Minimum	68.7907	75.7572
1st quartile	84.2624	84.8486
Median	87.2162	87.503
Mean	86.5376	86.8988
3rd quartile	89.5262	89.6185
Maximum	93.2732	93.0413
Range	24.4825	17.2841
Interquartile range	5.2638	4.7699
Variance	13.8434	12.3059
Standard deviation	3.7207	3.508
Skewness	−0.8964	−0.8003
Kurtosis	0.6558	0.2731

5. Case Study Application

5.1. Exploratory Data Analysis

A statistical summary of the geochemical attributes for the 10% sample is shown in Table 2.

Table 2. Statistics summary for a 10% sample for all geochemical attributes, where the numbering corresponds to: 1—clays, 2—chalcocite, 3—bornite, 4—chalcopyrite, 5—tennantite, 6—molybdenite, 7—pyrite, 8—copper (Cu), 9—molybdenum (Mo), and 10—arsenic (As), 11—copper recovery and 12—bond work index.

Statistics	1	2	3	4	5	6	7	8	9	10	11	12
Size	347	347	347	347	347	347	347	347	347	347	347	347
Minimum	0.7294	0.0045	0.0047	0.205	0.0046	0.0047	0.0053	0.0766	0.0028	0.0009	75.7572	11.39
1st quartile	2.1378	0.009	0.0674	0.5892	0.005	0.0089	0.6665	0.3199	0.0054	0.001	84.8486	12.5312
Median	3.2251	0.0457	0.1308	0.789	0.0067	0.0127	1.6151	0.4129	0.0076	0.0014	87.503	12.8145
Mean	4.08	0.0804	0.1733	0.9003	0.0128	0.0185	1.8393	0.4569	0.0111	0.0026	86.8988	12.9392
3rd quartile	5.0217	0.1261	0.2246	1.0561	0.0123	0.0218	2.7769	0.5312	0.0131	0.0025	89.6185	13.1052
Maximum	20.7767	0.7682	0.9131	3.6683	0.133	0.125	7.1561	1.6023	0.075	0.027	93.0413	22.2781
Range	20.0473	0.7637	0.9084	3.4633	0.1285	0.1203	7.1509	1.5257	0.0722	0.0261	17.2841	10.8881
Interquartile range	2.884	0.1172	0.1571	0.4669	0.0072	0.0129	2.1103	0.2113	0.0077	0.0015	4.7699	0.5739
Variance	8.0645	0.0101	0.0255	0.2463	0.0003	0.0002	1.911	0.0461	0.0001	0	12.3059	0.9168
Standard deviation	2.8398	0.1003	0.1597	0.4963	0.0174	0.0153	1.3824	0.2146	0.0092	0.0035	3.508	0.9575
Skewness	1.7563	2.9569	1.7926	2.3118	4.4297	2.4914	0.664	1.7473	2.4914	4.4297	−0.8003	5.058
Kurtosis	4.3005	14.3695	3.6934	7.8782	23.0514	9.0065	−0.1015	4.9835	9.0065	23.0514	0.2731	36.8668

5.2. Variable Selection

A bivariate analysis is carried out to select the geochemical attribute that has a greater dependence relationship with respect to copper recovery. For this purpose, heat maps of the Pearson, Kendall, and Spearman correlation coefficients are obtained (see the latter in Figure 5).

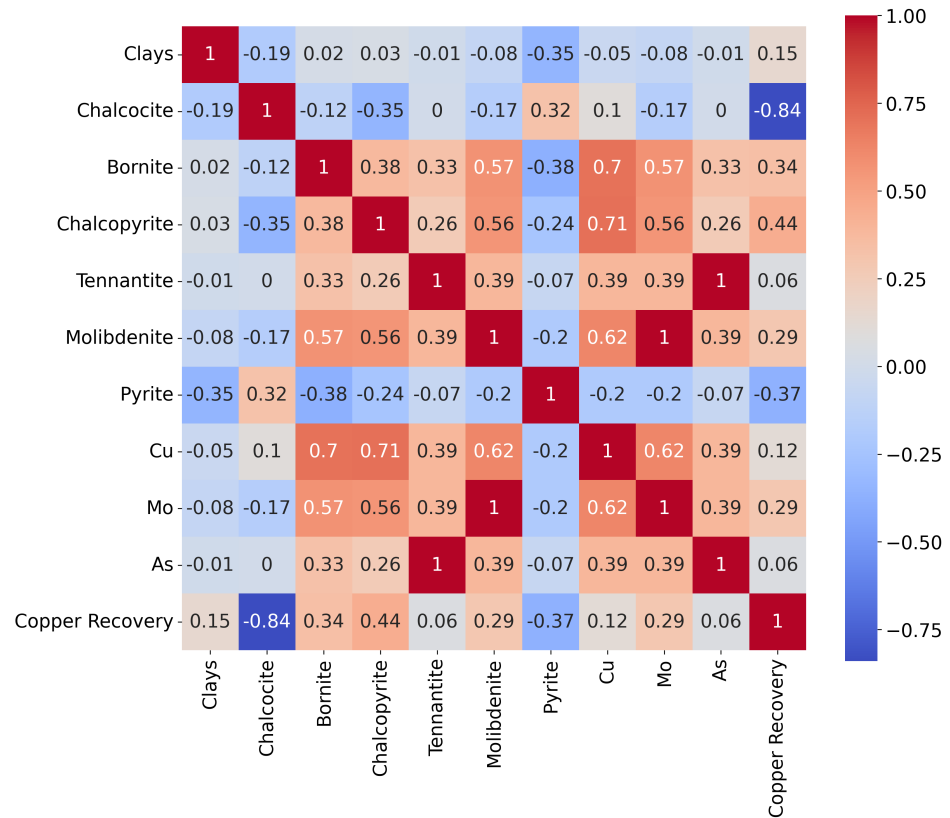


Figure 5. Spearman correlation heat map of all geochemical attributes for a 10% sample.

The geochemical attribute that shows the strongest correlation with copper recovery is chalcocite, with correlation coefficients of -0.75 , -0.84 , and -0.63 according to Pearson, Spearman, and Kendall, respectively (see Table 3). However, it is important to note that the dependency relationship between the variables, as shown in Figure 6, presents a clearly nonlinear behavior. The variable chalcocite, which is mostly sampled, is selected to perform the comparison between the collocated cokriging and conditional quantile regression methods.

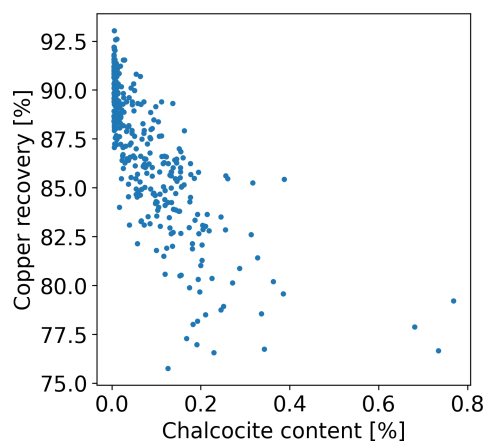


Figure 6. Copper recovery vs. chalcocite scatterplot for 10% sample dataset.

Table 3. Copper recovery and chalcocite correlation coefficients.

Primary Variable	Secondary Variable	Pearson	Kendall	Spearman
Copper recovery	chalcocite	−0.75	−0.63	−0.84

5.3. Modeling

As was mentioned in Section 3, the modeling step is different for each method. CCM corresponds to a *spatial correlation modeling* step, while CQRM to a *dependence modeling* step, as shown below.

5.3.1. Spatial Correlation Modeling

The spatial correlation model for CCM consists of estimating the sample variogram and optimally adjusting a variogram model to the primary variable, which in this case is copper recovery.

The chalcocite spatial distribution for the complete dataset on the left side and the corresponding map for the 10% sampling on the right side are shown in Figure 7, while Figure 8 shows the copper recovery empirical variogram calculated by Matheron estimator and a spherical fitted model with an effective range of 141.12 m, a sill of 12.20, and no nugget effect. This spatial correlation model is used in conjunction with Pearson’s correlation coefficient in the collocated cokriging application to chalcocite complete dataset (left side of Figure 7).

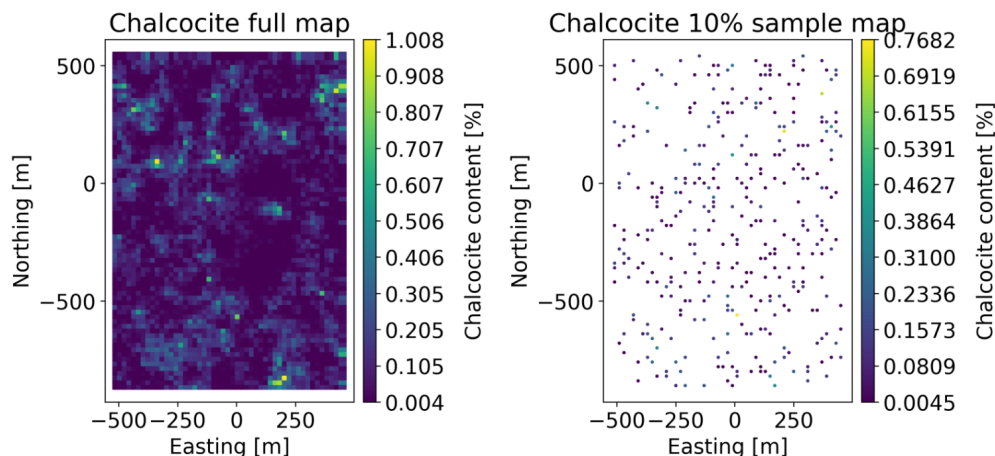


Figure 7. Chalcocite for the 100% (in the left) and 10% (in the right) sample maps, respectively.

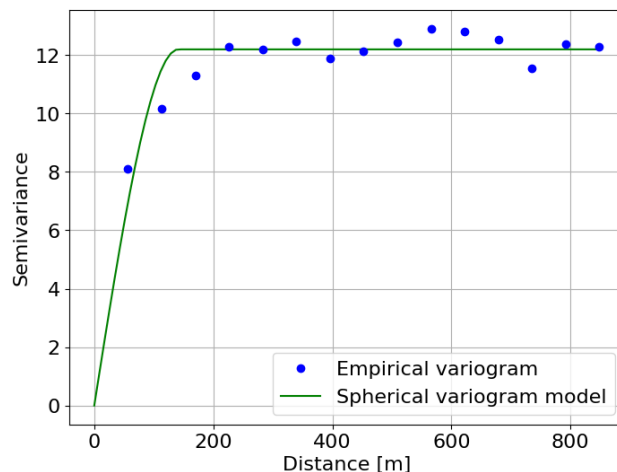


Figure 8. Copper recovery variogram model. The blue dots are the empirical variogram and the continuous green line is the fitted variogram model.

5.3.2. Dependence Modeling

The distribution of chalcocite shows a strong positive skewness (see Figure 9), which probably will not fit a parametric distribution, while copper recovery presents a slight negative skewness (see Figure 10).

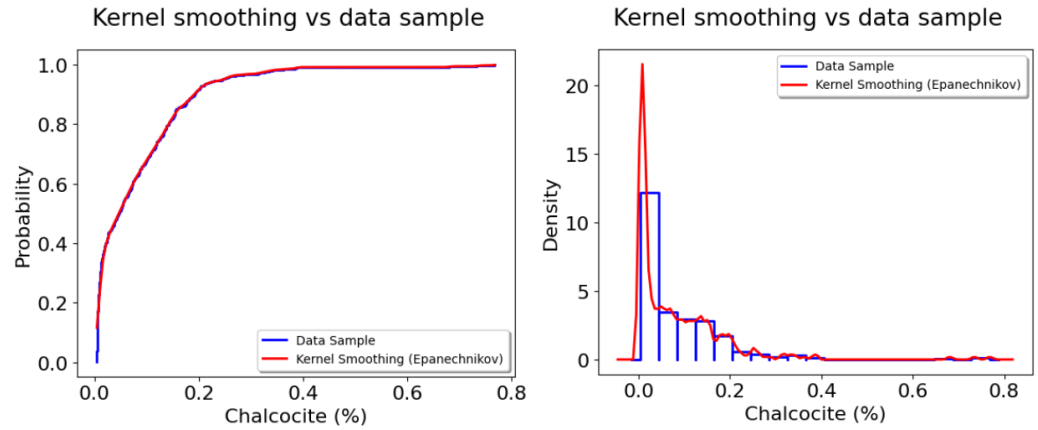


Figure 9. Model fitting of the chalcocite marginal using the kernel smoothing method with the Student function. The cumulative distribution function is on the left side and the probability density function is on the right side.

For both chalcocite and copper recovery, marginal distributions are fitted using the nonparametric method of kernel smoothing, which estimates a probability density function (PDF) from data without assuming a specific form. In particular, an Epanechnikov kernel [40] is the best fitting function for both marginals using the Bayesian information criterion (see Figures 9 and 10). Subsequently, a copula model is fitted using the same kernel smoothing method and criterion, but the best kernel is a Student function [41], which shows a better fit than other evaluated kernels, including the Epanechnikov one (see Figure 11).

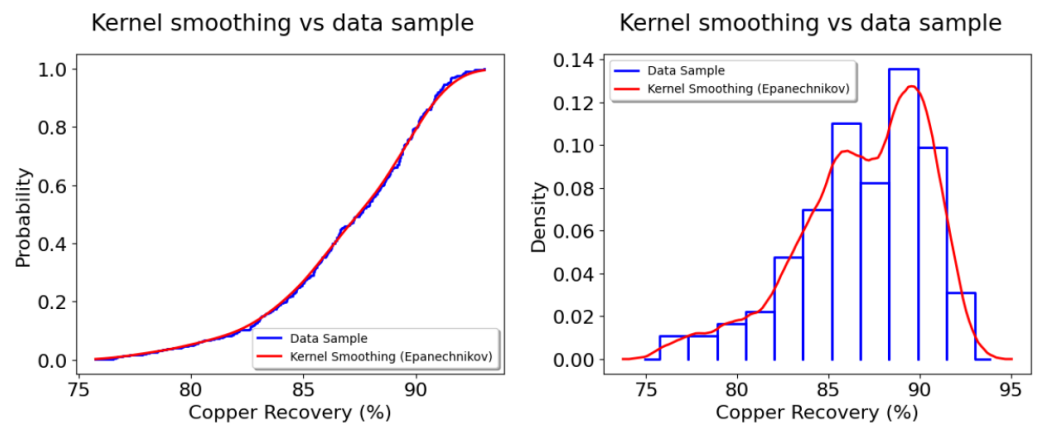


Figure 10. Model fitting of the copper recovery marginal using the kernel smoothing method with the Student function. The cumulative distribution function is on the left side and the probability density function is on the right side.

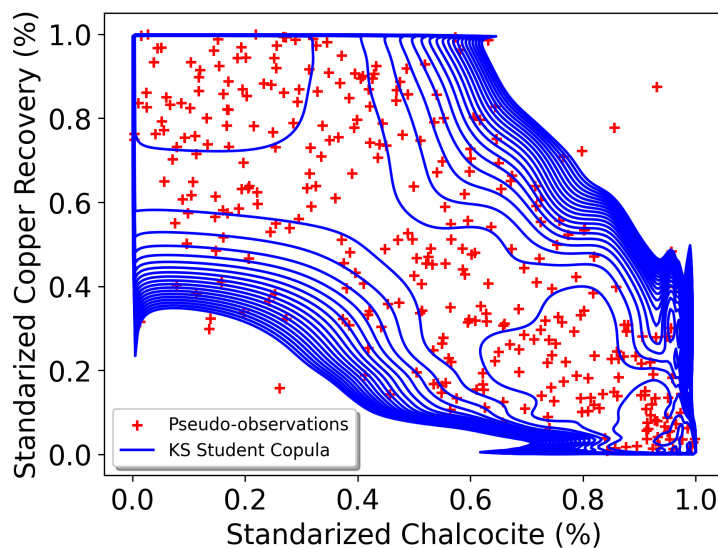


Figure 11. Copula model fitting using kernel smoothing method with a kernel Student function.

5.4. Validation

The validation stage consists of two parts. The first part involves verifying that the dependency model accurately reproduces both the univariate statistics of the marginal variables (chalcocite and copper recovery) and the bivariate statistics of the data sample. This is achieved through a joint nonconditional simulation using the dependency model (Appendix A.3). The second part assesses the predictive power of the quantile regression method by comparing its results with the 10% data sample (Appendix A.4). Specifically, copper recovery values are estimated by applying the quantile regression method conditioned on the chalcocite data from the 10% sample and are then compared with the actual copper recovery values from that sample.

The nonconditional simulation is generated using the bivariate probability distribution function of chalcocite and copper recovery from the previous section. To ensure comparability, the unconditional simulation produces a sample of the same size as the 10% data sample. The resulting simulation is statistically equivalent to the 10% data sample (see Tables 4 and 5). Additionally, the dependency model accurately reproduces the dependence pattern between the chalcocite data and copper recovery (see Figure 12), thereby validating both the models of the marginal distributions and the copula model.

Table 4. Statistics summary for copper recovery and chalcocite nonconditional joint simulation.

Statistics	Chalcocite	Copper Recovery
Size	347	347
Minimum	−0.0051	75.8326
1st quartile	0.0109	84.4898
Median	0.0540	87.522
Mean	0.0872	86.8948
3rd quartile	0.1440	89.7867
Maximum	0.7676	93.2552
Range	0.7727	17.4227
Interquartile range	0.1331	5.2969
Variance	0.0108	13.7954
Standard deviation	0.1040	3.7142
Skewness	2.5942	−0.6636
Kurtosis	14.6298	2.8307

Table 5. Copper recovery and chalcocite nonconditional simulation correlation coefficients.

Primary Variable	Secondary Variable	Pearson	Kendall	Spearman
Copper recovery	Chalcocite	−0.69	−0.52	−0.73

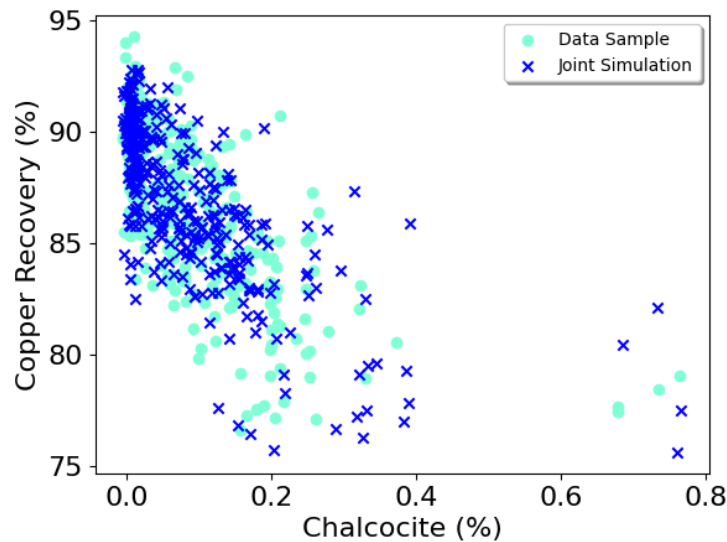


Figure 12. Joint copper recovery and chalcocite nonconditional simulation. The aquamarine circles are the data sample values and blue crosses are the copper–chalcocite joint bivariate unconditional simulation values.

Conditional quantile regression is applied to predict copper recovery conditional on chalcocite data values. These results are compared to actual copper recovery values, which are not used in the modeling. The estimated values of the median and the first (Q1) and third (Q3) quartiles (50% quantile, 25% quantile and 75% quantile, respectively) by CQRM follow the global trend of the real data (Figure 13) and their spatial distribution.

When comparing the omnidirectional variogram obtained from the copper recovery of the complete dataset with the copper recovery variogram estimated through the CQRM, it is observed that both are quite close (see Figure 14). This finding demonstrates the excellent ability to reproduce spatial variability using CQRM at unsampled locations.

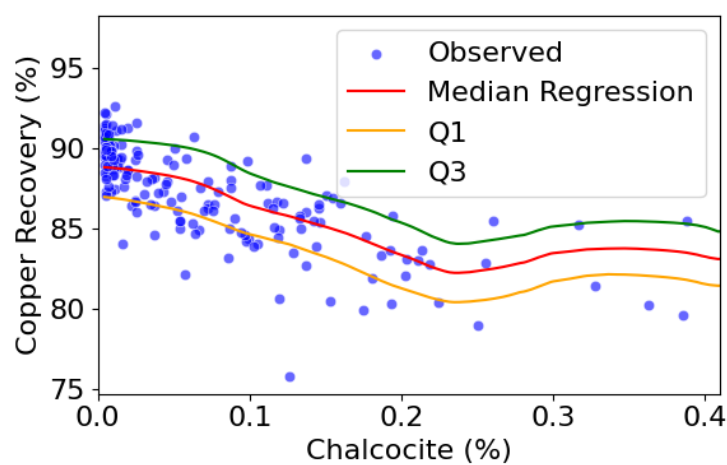


Figure 13. Scatterplot of copper recovery vs. chalcocite. The blue dots are the observed data values, the red line is the median estimated values, and the orange and green lines are the first (Q1) and third (Q3) quartiles, respectively, by CQRM application.

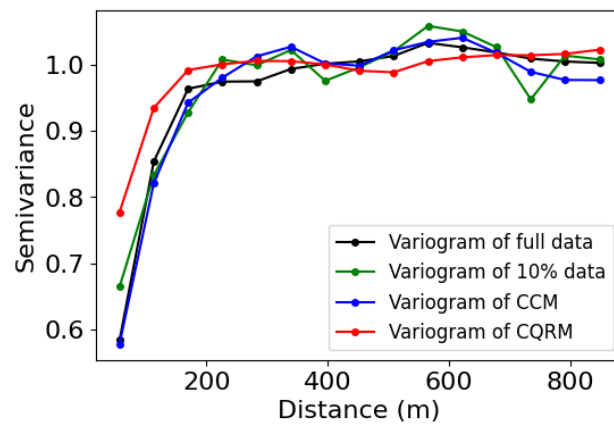


Figure 14. Copper recovery empirical variograms from full data, 10% sample, CCM and CQRM.

5.5. Prediction

The collocated cokriging method (CCM) was applied with the Markov 1 model, described in detail in [42], where chalcocite was employed as an exhaustive secondary variable. The co-estimation plan considered a search neighborhood equivalent to the effective range of the spherical variogram, at 145 radial meters, using a limit of 20 observations as the maximum for each co-estimation. Similarly, the conditional quantile regression method (CQRM) was applied to each chalcocite cell using the dependence model obtained in Section 5.3.2 for the estimation of the median and interquartile range of copper recovery.

The results of the application of both methods are shown numerically in terms of performance metrics in Table 6 and graphically in Figures 15 and 16.

Table 6. Performance metrics of collocated cokriging vs. quantile regression method for copper recovery estimation conditioned by chalcocite for 10% sample data, where RMSE is the root mean squared error, MAE is the mean absolute error, MAPE is the mean absolute percentage error, and R² is the determination coefficient.

Performance Metrics	CCM	CQRM
RMSE	3.85	2.86
MAE	3.03	1.71
MAPE	3.58	2.04
R ²	0.09	0.67

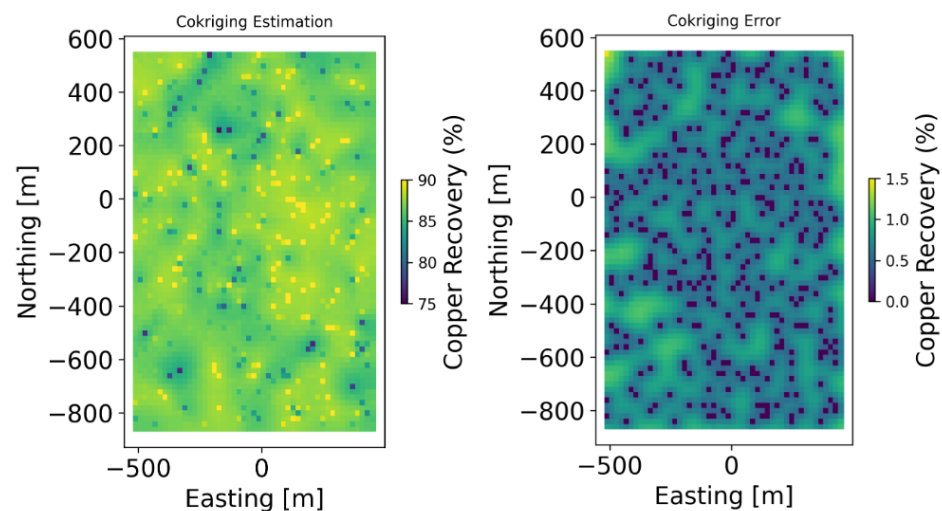


Figure 15. Mean estimation (on the left side) and standard deviation (on the right side) for copper recovery conditioned by chalcocite using CCM.

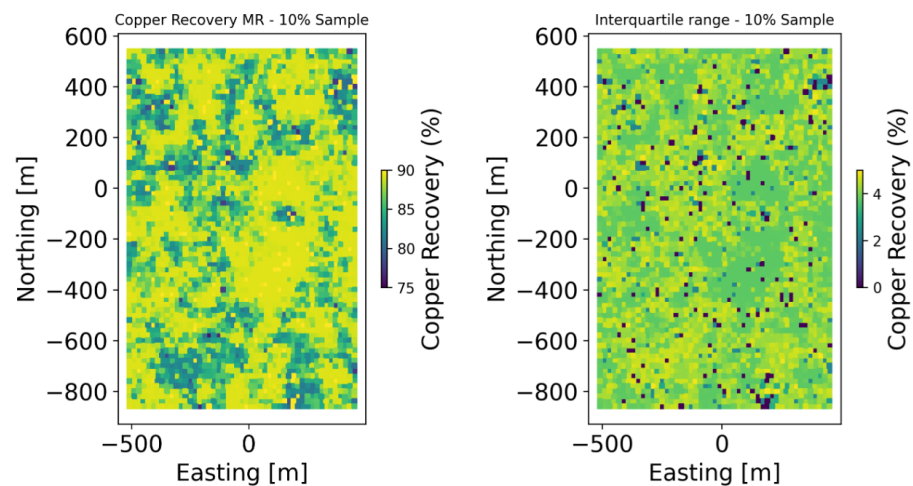


Figure 16. Median estimation (on the left side) and interquartile range (on the right side) for copper recovery conditioned by chalcocite using CQRM.

The performance metrics (RME, MAE, and MAPE) in Table 6 are significantly lower for CQRM relative to CCM, meaning that CQRM is a much more accurate method, and R^2 corroborates that CQRM has a better goodness of fit. Or, in other words, it explains a greater percentage of the variance in the data than the CCM. This is complemented graphically, since it can be seen that the spatial distribution pattern of the copper recovery estimate using CCM (Figure 15) is very different from that of Figure 2 of the dataset complete, while the CQRM estimation (Figure 16) has a much closer spatial distribution.

Performance Evaluation

By reducing heterotopy through an increase in the sampling of the primary variable, a substantial decrease in error is observed. This improvement is logical and is due to the fact that a greater amount of data for the primary variable allows for a better understanding of its behavior and its relationship with other variables.

A comparison of CQRM performance under different scenarios is shown in Figure 17. In particular, five scenarios were evaluated with a considerably high heterotopia, with different percentages (1%, 5%, 10%, 15%, and 20%) of the primary copper recovery variable, but preserving the set of complete data for the secondary variable chalcocite. It can be seen that with more information, the CQRM can more accurately estimate the primary copper recovery variable, resulting in a significant reduction in uncertainty and, therefore, the associated error. This is clearly demonstrated in Table 7, where the performance metrics consistently improve as the percentage of data for the primary variable increases.

Table 7. Performance metrics of quantile regression method for copper recovery estimation conditioned by chalcocite under different scenarios: 1%, 5%, 10%, 15%, and 20% sample data, where RMSE is the root mean squared error, MAE is the mean absolute error, MAPE is the mean absolute percentage error, and R^2 is the determination coefficient.

Performance Metrics	1%	5%	10%	15%	20%
RMSE	3.78	2.97	2.86	2.35	2.24
MAE	2.64	1.92	1.71	1.52	1.41
MAPE	3.16	2.29	2.04	1.81	1.68
R^2	0.39	0.63	0.67	0.79	0.81

The results are notable considering the simplicity of this approach, as it depends only on a single secondary variable, and the computational efficiency it presents.

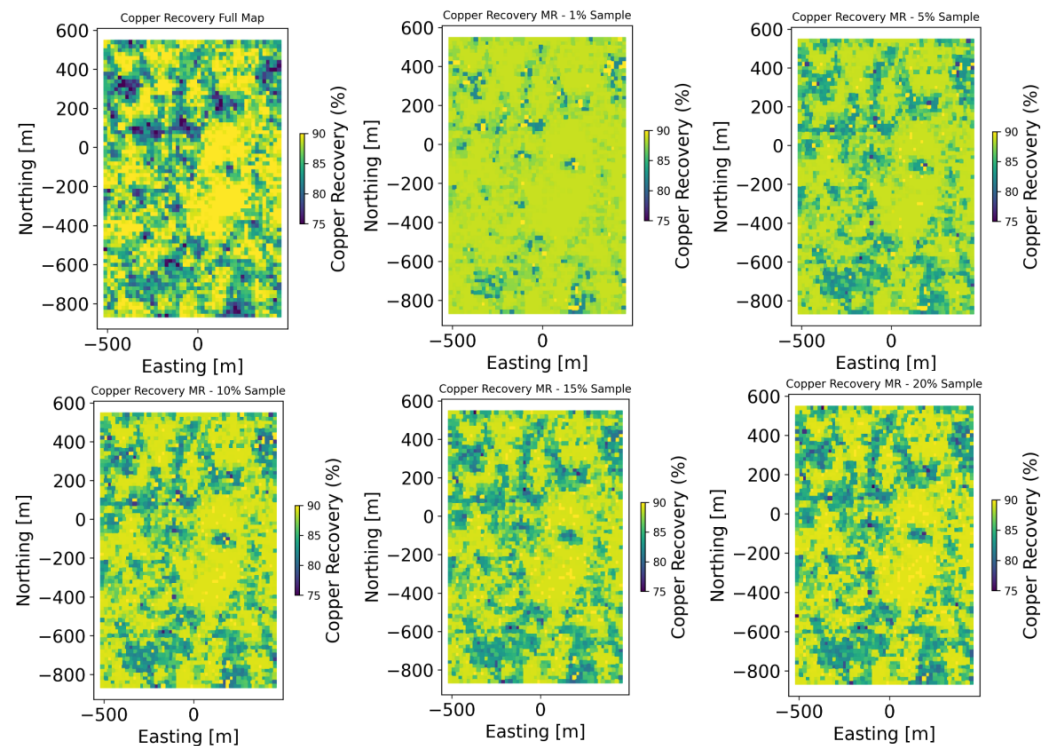


Figure 17. Median quantile regression estimation for copper recovery conditioned by chalcocite under different scenarios. The figures are arranged from left to right and from top to bottom: full (100%) map, 1%, 5%, 10%, 15%, and 20% sample data.

6. Discussion

The results of the CCM, under the same heterotopic sampling conditions of 10%, did not surpass those of the CQRM, even when the latter was used under less favorable sampling conditions (Table 7). In particular, it can be observed that the performance metrics for the 1% sample data are even better for CQRM than those for the 10% sample data for CCM. This is because methods like the CCM depend on strictly linear relationships between variables and are not competitive in this case, where there is no linearity and the dependencies are complex.

Although the CQRM does not explicitly incorporate a spatial correlation model for copper metallurgical recovery, it reproduces its variability quite well. When observing Figure 14, there is a high similarity between the data sample and prediction variograms. This occurs because the primary variable (copper recovery) inherits the spatial correlation through the secondary variable using the joint dependence model. By not requiring a spatial correlation model, a certain degree of subjectivity is eliminated from the process, adding practicality and efficiency. It is shown that a copula-based dependence model can reproduce spatial variability, provided that the joint dependence model is established following the proposed methodological guidelines. This includes appropriately selecting the predictor variable based on maximum dependence, correctly modeling marginal distributions, and choosing an appropriate bivariate model, as well as validating the model through simulation to reproduce the dependence of the sampling data.

7. Conclusions

Copper metallurgical recovery is a key variable in defining the cutoff grade for mining operations, directly influencing the amount of copper that can be profitably extracted and processed. Consequently, it is critical in the economic evaluation of a project. In this context, an accurate estimation of metallurgical recovery is vital for strategic planning and decision making.

The approach proposed in this article consists of the application of a novel conditional quantile regression method for metallurgical copper recovery conditioned on a secondary variable that is widely sampled and exhibits the maximum possible dependence. The particularity of this approach lies in the optimal estimation of the codependency model of the variables using a kernel smoothing method for the marginals as well as for the copula.

The presented application is based on a synthetic but realistic case of a porphyry copper deposit, where the median regression for copper recovery is estimated at all known chalcocite locations, used as an explanatory variable. The comparison of the results obtained by CQRM versus CCM shows that, both in terms of precision and efficiency, as well as reproduction of spatial variability, the method proposed in this article is superior when the dependencies are nonlinear.

Although this is a specific application to a case study, the methodology can be extended to other scenarios, such as an exploration drilling campaign, where the largest scale of information comes from drilling interval logs, with metallurgical recovery measured in fewer quantities and not necessarily in the same locations.

CQRM is an innovative, practical, efficient, and versatile approach. It avoids strong assumptions about the data and theoretical constraints in its implementation, demonstrating a clear capacity to adapt to different levels of information and scales, making it applicable to a wide variety of scenarios.

As future work, the implementation of conditional quantile regression method based on copula model with multivariate dependencies is proposed. This model extension is expected to further improve the accuracy of predictions while maintaining practicality, versatility, and computational efficiency.

As can be deduced from the Funding section of this manuscript, this research has raised the interest of other sectors, such as the steel industry, who wishes to maximize the efficiency of all its processes in terms of both industrial and social impact.

Author Contributions: Conceptualization, H.H. and M.A.D.-V.; methodology, M.A.D.-V.; software, M.A.D.-V.; validation, H.H.; formal analysis, M.A.D.-V.; writing—original draft preparation, H.H. and M.A.D.-V.; writing—review and editing, E.A., A.G., and A.O.-Z.; supervision, E.A.; funding acquisition, A.G. and A.O.-Z. All authors have read and agreed to the published version of the manuscript.

Funding: Work funded by project SILENCE—European Commission—Research Program of the Research Funds for Coal and Steel—Prj. No.: 101112516.

Data Availability Statement: Data used in this article are taken from the publication [19] and are publicly available.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CCM	Collocated cokriging method
CQRM	Conditional quantile regression method based on copulas
MAE	Mean absolute error
MAPE	Mean absolute percentage error
PDF	Probability density function
R ²	Determination coefficient
RMSE	Root mean squared error

Appendix A. Conditional Quantile Regression Method Based on Copulas

Sklar in 1959 [33] established a theorem indicating a functional relationship between the joint probability distribution function of a random vector and its univariate marginal distribution functions. For instance, in the case of two variables, if (X, Y) is a random vector with a joint probability distribution $H_{XY}(x, y) = P(X \leq x, Y \leq y)$, then the marginal

distribution functions of X and Y are $F_X(x) = P(X \leq x) = H_{XY}(x, \infty)$ and $G_Y(y) = P(Y \leq y) = H_{XY}(\infty, y)$, respectively. However, when marginalizing H_{XY} , some information is lost, as the marginal distributions F_X and G_Y generally do not suffice to fully determine H_{XY} . This is because the marginal distributions only describe the individual probability behavior of the random variables they represent. Sklar’s theorem demonstrates that there exists a function $C_{XY}: [0, 1]^2 \mapsto [0, 1]$ such that

$$H_{XY}(x, y) = C_{XY}(F_X(x), G_Y(y)) \tag{A1}$$

where C_{XY} is the *copula* function associated with the bivariate random vector (X, Y) and describes its dependence relationship, H is the bivariate probability distribution function, and F and G are the univariate (marginal) probability distributions.

Copula functions are a valuable tool for constructing joint probability models with greater flexibility. They allow us to independently select univariate models for the random variables of interest and choose a copula function that best represents their dependence, either parametrically or nonparametrically. For example, in a multivariate normal model, all marginal distributions must be normal, with no tail dependence and finite second moments for well-defined correlations. The multivariate normal model is a specific instance where the underlying copula is Gaussian and all univariate marginals follow a normal distribution.

When F_X and G_Y are continuous, elementary probability theory tells us that $U = F_X(X)$ and $V = G_Y(Y)$ are continuous uniform random variables on $(0, 1)$. The underlying copula C for the random vector (U, V) is the same copula corresponding to (X, Y) . According to Sklar’s theorem, the joint probability distribution function for (U, V) is given by $H_{UV}(u, v) = C(F_U(u), G_V(v)) = C(u, v)$. Therefore, if F_X and G_Y are known but H_{XY} is unknown, and we have an observed random sample $\{(x_1, y_1), \dots, (x_n, y_n)\}$ of (X, Y) , the set $\{(u_k, v_k) = (F_X(x_k), G_Y(y_k)) : k = 1, \dots, n\}$ will be an observed random sample of (U, V) with the same underlying copula C as (X, Y) . Since $C = F_{UV}$, we can use the values (u_k, v_k) (known as copula observations) to estimate C as a joint empirical distribution:

$$\hat{C}(u, v) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}_{\{u_k \leq u, v_k \leq v\}} \tag{A2}$$

Strictly, the estimate \hat{C} is not a copula since it is discontinuous and copulas are always continuous. If F_X, G_Y , and H_{XY} are all unknown, which is the most common case, F_X and G_Y are estimated by their empirical univariate distribution functions:

$$\hat{F}_X(x) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}\{x_k \leq x\}, \quad \hat{G}_Y(y) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}\{y_k \leq y\} \tag{A3}$$

where \mathbb{I} represents an indicator function equal to 1 when its argument is true and 0 otherwise.

We will refer to the set of pairs $\{(\hat{u}_k, \hat{v}_k) = (\hat{F}_X(x_k), \hat{G}_Y(y_k)) : k = 1, \dots, n\}$ as pseudo observations of the copula. It can be verified directly that $\hat{F}_X(x_k) = \frac{1}{n} \text{rank}(x_k)$ and $\hat{G}_Y(y_k) = \frac{1}{n} \text{rank}(y_k)$. In this case, the concept of empirical copula, see [34], is defined as the following function $C_n: I_n^2 \mapsto [0, 1]$, where $I_n = \left\{ \frac{i}{n} : i = 0, \dots, n \right\}$, given by

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}\{\text{rank}(x_k) \leq i, \text{rank}(y_k) \leq j\} \tag{A4}$$

Again, C_n is not a copula but is an estimate of the underlying copula in the mesh I_n^2 that can be extended to a copula in $[0, 1]^2$ of, for example, Bernstein polynomials, as proposed

and studied in [43], leading to what is known as a nonparametric estimate of the Bernstein copula $\tilde{C}: [0, 1]^2 \mapsto [0, 1]$ given by

$$\tilde{C}(u, v) = \sum_{i=0}^n \sum_{j=0}^n C_n \left(\frac{i}{n}, \frac{j}{n} \right) \binom{n}{i} u^i (1-u)^{n-i} \binom{n}{j} v^j (1-v)^{n-j} \tag{A5}$$

Appendix A.1. Three Approaches to Building a Copula-Based Dependency Model

There are three approaches to building a copula-based dependency model: parametric, nonparametric, and semiparametric.

The parametric approach consists of being able to fit a known copula model to the empirical copula, such as the Frank, Gumbel, or Clayton copulas, which belong to the family of Archimedean copulas [34,35], as well as being able to fit the empirical marginal probability distributions to known distribution functions such as normal or Gaussian, lognormal, gamma, Weibull, etc. In this way, a joint probability distribution model is obtained.

The nonparametric approach consists of numerically approximating the empirical copula and its marginals, usually by means of some polynomial expression. In this approach, Bernstein polynomials [28,43] and splines [44], as well as the kernel smoothing method, have been used [35], where kernel smoothing is a type of weighted moving average.

While a semiparametric approach is a combination of the two previous approaches, this allows two options, that is, a model could be fitted to the empirical copula and approximate the marginals with a polynomial, or the empirical copula could be approximated by a polynomial expression and adjust the marginals with a known distribution model [45].

Appendix A.2. Kernel Density Estimation with Kernel Smoothing Method

Consider (x_1, x_2, \dots, x_n) as independent and identically distributed samples drawn from a univariate distribution characterized by an unknown density function f at any arbitrary point x . Our focus lies in approximating the shape of this function f . Its kernel density estimator is expressed as follows:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \tag{A6}$$

In this context, K denotes the kernel, a function that yields only non-negative values, while $h > 0$ represents a smoothing parameter referred to as the bandwidth. A kernel labeled with the subscript h is termed the scaled kernel, defined by $K_h(x) = \frac{1}{h}K(\frac{x}{h})$. In essence, one aims to select h to be as small as the data permits intuitively; nevertheless, there invariably exists a trade-off between the estimator’s bias and its variance.

Appendix A.3. Copula Simulation Algorithm

As summarized in [45], in order to simulate the repetitions of the random vector (X, Y) with the dependence structure inferred from the observed data $(x_1, y_1), \dots, (x_n, y_n)$, we have the following algorithm:

- (i) Generate two independent and continuous random variables u and t uniformly distributed in $(0, 1)$.
- (ii) Set $v = c_u^{-1}(t)$, where $c_u(v) = \frac{\partial \tilde{C}(u,v)}{\partial u}$.
- (iii) The desired pair is $(x, y) = (\tilde{Q}_n(u), \tilde{R}_n(v))$, where \tilde{Q}_n and \tilde{R}_n are the empirical quantile functions for X and Y , respectively.

Appendix A.4. Conditional Quantile Regression Algorithm

For a value x in the range of the random variable X and a given $0 < \alpha < 1$, let $y = \varphi_\alpha(x)$ be the solution of the equation $P(Y \leq y|X = x) = \alpha$. Then the graph of

$y = \varphi_\alpha(x)$ is the *quantile regression curve* α of Y conditional on $X = x$. In [34] it is proven that

$$P(Y \leq y | X = x) = c_u(v)|_{u=F_X(x),v=G_Y(y)} \tag{A7}$$

This result leads to the following algorithm to obtain *quantile regression curve* α of Y conditional on $X = x$:

(i) Set $c_u(v) = \alpha$.

(ii) Solve the regression curve for v :

$$v = g_\alpha(u). \tag{A8}$$

(iii) Replace u by $\tilde{Q}_n^{-1}(x)$ and v by $\tilde{R}_n^{-1}(y)$.

(iv) Solve the regression curve for y :

$$y = \varphi_\alpha(x). \tag{A9}$$

Appendix B. Collocated Cokriging Method with Markov Model

The kriging method is well known as the best unbiased linear spatial estimator of a single random function and the cokriging method is its generalization for two or more random functions. This requires calculating the primary variogram (e.g., copper recovery), the secondary variogram (e.g., geochemical attribute), and the cross-variogram [38,46]. In particular, the cross-variogram takes into account the spatial dependence between the two random functions and is defined as follows:

$$\gamma_{12}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^N (Z_1() - Z_1(x_\alpha + h))(Z_2(x_\alpha) - Z_2(x_\alpha + h)) \tag{A10}$$

where γ_{12} is the semivariance between the random functions Z_1 (copper recovery) and Z_2 (geochemical attribute), h is the lag distance, x_α is a spatial location, and N is the number of lag distances.

The rigorous application of the ordinary cokriging method is frequently very difficult and complicated since it also requires that the set of variograms comply with the linear coregionalization model. This model is very restrictive since it requires that all variograms fit the same model with a common range.

A more efficient and practical alternative is the collocated cokriging method that was introduced by Almeida and Journel (1994) [47]. It is a variant of cokriging method for spatial interpolation that leverages both primary and secondary data. But it simplifies the computation by using only the secondary variable's value that is collocated (at the same location) as the primary variable. This method is particularly useful when secondary data are more densely sampled than primary data.

However, this approach requires a conditional independence between the primary variable and the secondary variable, given their collocated values. This is known as a Markov assumption of conditional independence, which simplifies the modeling process. There are two types of Markov models: Markov models 1 and 2, respectively. Here, Markov model 1 (MM1) is used as it is the simplest and most straightforward option.

In Markov model 1, the following conditional independence assumption is made:

$$E(Z_2(u) | Z_1(u) = z_1, Z_1(u') = z_1(u')) = E(Z_2(u) | Z_1(u) = z_1) \tag{A11}$$

This implies, under the assumptions of MM1, a simplification in the statistical relationship between the primary and secondary variables. Thus, the cross-correlogram model can be written as

$$\rho_{12}(h) = \rho_{12}(0)\rho_1(h) \tag{A12}$$

where the correlogram ρ is expressed by

$$\rho(h) = 1 - \gamma(h) \tag{A13}$$

Assuming that $\gamma(h)$ is the variogram for data with a standard normal distribution, the correlogram is simply the variogram inverted and shifted upward by one. While the variogram measures spatial covariance, the correlogram measures spatial correlation. Therefore, $\rho_{12}(0)$, the correlogram at a lag distance of zero, is equivalent to the correlation coefficient between variables 1 and 2.

In summary, to apply the collocated cokriging method with Markov model 1 (MM1), all that is needed is the primary variable variogram and the linear correlation coefficient between the primary and secondary variables. The variogram for secondary variable is not necessary.

Note that you MUST perform a normal score transformation for the primary and secondary data prior to invoking MM1.

References

1. Kasa, F.K.; Dağ, A. Appraising economic uncertainty in open-pit mining based on fixed and variable metallurgical recovery. *Arch. Min. Sci.* **2022**, *67*, 699–713.
2. Gholami, A.; Asgari, K.; Khoshdast, H.; Hassanzadeh, A. A hybrid geometallurgical study using coupled Historical Data (HD) and Deep Learning (DL) techniques on a copper ore mine. *Physicochem. Probl. Miner. Process.* **2022**, *58*, 147841. [[CrossRef](#)]
3. Oumesaoud, H.; Faouzi, R.; Aboulhassan, M.A.; Naji, K.; Benzakour, I.; Faqir, H.; Oukhrib, R.; Elboughdiri, N. Iron Oxide–copper Mineral Associations in Supergene Zones: Insights into Flotation Challenges and Optimization Using Response Surface Methodology. *ACS Omega* **2024**, *9*, 24438–24452. [[CrossRef](#)]
4. Madenova, Y.; Madani, N. Application of Gaussian Mixture Model and Geostatistical Co-simulation for Resource Modeling of Geometallurgical Variables. *Nat. Resour. Res.* **2021**, *30*, 1199–1228. [[CrossRef](#)]
5. Ramos Armijos, N.J.; Calderón Celis, J.M. Revisión del modelo geometalúrgico para la estimación de recursos minerales en depósitos pórfido cupríferos. *Rev. Del Inst. De Investig. De La Fac. De Minas Metal. Y Cienc. Geográficas* **2022**, *25*, 445–459. [[CrossRef](#)]
6. Mu, Y.; Salas, J.C. Data-Driven Synthesis of a Geometallurgical Model for a copper Deposit. *Processes* **2023**, *11*, 1775. [[CrossRef](#)]
7. Hunt, J.A.; Berry, R.F. Economic Geology Models #3. Geological Contributions to Geometallurgy: A Review. *Geosci. Can.* **2017**, *44*, 103–118. [[CrossRef](#)]
8. Hoffmann, J.; Augusto, J.; Resende, L.; Mathias, M.; Mazzinghy, D.; Bianchetti, M.; Mendes, M.; Souza, T.; Andrade, V.; Domingues, T.; et al. Modeling Geospatial Uncertainty of Geometallurgical Variables with Bayesian Models and Hilbert-Kriging. *Math. Geosci.* **2022**, *54*, 1227–1253. [[CrossRef](#)]
9. Hunt, J.; Kojovic, T.; Berry, R. Estimating comminution indices from ore mineralogy, chemistry and drill core logging. In Proceedings of the Second AusIMM International Geometallurgy Conference (GeoMet), Carlton, VIC, Australia, 30 September–2 October 2013; Dominy, S., Ed.; The Australasian Institute of Mining and Metallurgy (AusIMM): Carlton, Australia, 2013; pp. 173–176.
10. Garrido, M.; Ortiz, J.; Sepúlveda, E.; Farfán, L.; Townley, B. An overview of good practices in the use of geometallurgy to support mining reserves in copper sulfides deposits. In Proceedings of the Conference: Procemin Geomet 2019, Santiago, Chile, 20–22 November 2019. [[CrossRef](#)]
11. Dowd, P.A.; Xu, C.; Coward, S. Strategic mine planning and design: Some challenges and strategies for addressing them. *Min. Technol.* **2016**, *125*, 22–34. [[CrossRef](#)]
12. Little, L.; Mclennan, Q.; Prinsloo, A.; Muchima, K.; Kaputula, B.; Siame, C. Relationship between ore mineralogy and copper recovery across different processing circuits at Kansanshi mine. *J. S. Afr. Inst. Min. Metall.* **2018**, *118*, 1155–1162. [[CrossRef](#)]
13. Adeli, A.; Dowd, P.; Emery, X.; Xu, C. Using cokriging to predict metal recovery accounting for non-additivity and preferential sampling designs. *Miner. Eng.* **2021**, *170*, 106923. [[CrossRef](#)]
14. Dachri, K.; Bouabidi, M.; Naji, K.; Nouar, K.; Benzakour, I.; Oummouch, A.; Hibti, M.; El Amari, K. Predictive insights for copper recovery: A synergistic approach integrating variability data and machine learning in the geometallurgical study of the Tizert deposit, Morocco. *J. Afr. Earth Sci.* **2024**, *212*, 105208. [[CrossRef](#)]
15. Lishchuk, V.; Lund, C.; Ghorbani, Y. Evaluation and comparison of different machine-learning methods to integrate sparse process data into a spatial model in geometallurgy. *Miner. Eng.* **2019**, *134*, 156–165. [[CrossRef](#)]
16. Chehreh Chelgani, S.; Nasiri, H.; Alidokht, M. Interpretable modeling of metallurgical responses for an industrial coal column flotation circuit by XGBoost and SHAP-A “conscious-lab” development. *Int. J. Min. Sci. Technol.* **2021**, *31*, 1135–1144. [[CrossRef](#)]
17. Cook, R.; Monyake, K.C.; Hayat, M.B.; Kumar, A.; Alagha, L. Prediction of flotation efficiency of metal sulfides using an original hybrid machine learning model. *Eng. Rep.* **2020**, *2*, e12167. [[CrossRef](#)]
18. Flores, V.; Leiva, C. A Comparative Study on Supervised Machine Learning Algorithms for copper Recovery Quality Prediction in a Leaching Process. *Sensors* **2021**, *21*, 2119. [[CrossRef](#)] [[PubMed](#)]
19. Garrido, M.; Sepúlveda, E.; Ortiz, J.; Townley, B. Simulation of synthetic exploration and geometallurgical database of porphyry copper deposits for educational purposes. *Nat. Resour. Res.* **2020**, *29*, 3527–3545. [[CrossRef](#)]

20. Campos, P.H.A.; Costa, J.F.C.L.; Koppe, V.C.; Bassani, M.A.A. Geometallurgy-oriented mine scheduling considering volume support and non-additivity. *Min. Technol.* **2022**, *131*, 1–11. [[CrossRef](#)]
21. Tavares, L.M.; Kalleback, R.D. Grindability of binary ore blends in ball mills. *Miner. Eng.* **2013**, *41*, 115–120. [[CrossRef](#)]
22. Carrasco, P.; Chilès, J.P.; Séguret, S.A. Additivity, metallurgical recovery, and grade. In Proceedings of the 8th International Geostatistics Congress, Santiago, Chile, 1–5 December 2008; pp. on CD. hal-00776943
23. Rossi, M.E.; Deutsch, C.V. *Mineral Resource Estimation*; Springer: Dordrecht, The Netherlands, 2014. [[CrossRef](#)]
24. Sohrabian, B.; Soltani-Mohammadi, S.; Pourmirzaee, R.; Carranza, E.J.M. Geostatistical Evaluation of a Porphyry Copper Deposit Using Copulas. *Minerals* **2023**, *13*, 732. [[CrossRef](#)]
25. Bárdossy, A.; Li, J. Geostatistical interpolation using copulas. *Water Resour. Res.* **2008**, *44*. [[CrossRef](#)]
26. Hernández-Maldonado, V.; Díaz-Viera, M.A.; Erdely, A. A joint stochastic simulation method using the Bernstein copula as a flexible tool for modeling nonlinear dependence structures between petrophysical properties. *J. Pet. Sci. Eng.* **2012**, *90–91*, 112–123. [[CrossRef](#)]
27. Kazianka, H.; Pilz, J. Spatial Interpolation Using Copula-Based Geostatistical Models. In *geoENV VII—Geostatistics for Environmental Applications*; Atkinson, P.M., Lloyd, C.D., Eds.; Springer: Dordrecht, The Netherlands, 2010; pp. 307–319. [[CrossRef](#)]
28. Díaz-Viera, M.A.; Erdely, A.; Kerdan, T.; del Valle-García, R.; Mendoza-Torres, F. Bernstein Copula-Based Spatial Stochastic Simulation of Petrophysical Properties Using Seismic Attributes as Secondary Variable. In *Geostatistics Valencia 2016. Quantitative Geology and Geostatistics, vol 19*; Gómez-Hernández, J., Rodrigo-Illari, J., Rodrigo-Clavero, M., Cassiraga, E., Vargas-Guzmán, J., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2017; pp. 487–504.
29. Le, V.H. Copula-Based Modeling for Petrophysical Property Prediction Using Seismic Attributes as Secondary Variables. Ph.D. Thesis, Instituto de Geofísica, UNAM, Mexico City, Mexico, 2021.
30. Kolev, N.; Paiva, D. Copula-based regression models: A survey. *J. Stat. Plan. Inference* **2009**, *139*, 3847–3856. [[CrossRef](#)]
31. Kim, J.M.; Cho, C.; Jun, C.; Kim, W.Y. The Changing Dynamics of Board Independence: A Copula Based Quantile Regression Approach. *J. Risk Financ. Manag.* **2020**, *13*, 254. [[CrossRef](#)]
32. Tepegozova, M.; Czado, C. Bivariate vine copula based regression, bivariate level and quantile curves. *arXiv* **2022**, arXiv:2205.02557.
33. Sklar, A. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **1959**, *8*, 229–231.
34. Nelsen, R.B. *An Introduction to Copulas*; Springer: New York, NY, USA, 2006. [[CrossRef](#)]
35. Joe, H. *Dependence Modeling with Copulas*; Chapman and Hall/CRC: Amsterdam, The Netherlands, 2014.
36. Tholana, T.; Musingwini, C. A Probabilistic Block Economic Value Calculation Method for Use in Slope Designs under Uncertainty. *Minerals* **2022**, *12*, 437. [[CrossRef](#)]
37. Jamshidi, M.; Osanloo, M. Determination of block economic value in multi-element deposits. In Proceedings of the 6th International Conference on Computer Applications in the Minerals Industries, CAMI2016-06, Istanbul, Turkey, 5–7 October 2016.
38. Wackernagel, H. Collocated Cokriging. In *Multivariate Geostatistics: An Introduction with Applications*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 165–169. [[CrossRef](#)]
39. Wand, M.; Jones, M. *Kernel Smoothing*; Chapman and Hall/CRC: New York, NY, USA, 1994.
40. Scott, D.W. *Multivariate Density Estimation: Theory, Practice, and Visualization*; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1992. [[CrossRef](#)]
41. Li, R.; Nadarajah, S. A review of Student's t distribution and its generalizations. *Empir. Econ.* **2020**, *58*, 1461–1490. [[CrossRef](#)]
42. Shmaryan, L.; Journel, A. Two Markov Models and Their Application. *Math. Geol.* **1999**, *31*, 965–988. [[CrossRef](#)]
43. Sancetta, A.; Satchell, S. The bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econom. Theory* **2004**, *20*, 535–562. [[CrossRef](#)]
44. Shen, X.; Zhu, Y.; Song, L. Linear B-spline copulas with applications to nonparametric estimation of copulas. *Comput. Stat. Data Anal.* **2008**, *52*, 3806–3819. [[CrossRef](#)]
45. Erdely, A.; Diaz-Viera, M. Nonparametric and Semiparametric Bivariate Modeling of Petrophysical Porosity-Permeability Dependence from Well Log Data. In *Copula Theory and Its Applications: Proceedings of the Workshop, Warsaw, Poland, 25–26 September 2009*; Jaworski, P., Durante, F., Härdle, W.K., Rychlik, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 267–278.
46. Babak, O.; Deutsch, C.V. Improved spatial modeling by merging multiple secondary data for intrinsic collocated cokriging. *J. Pet. Sci. Eng.* **2009**, *69*, 93–99. [[CrossRef](#)]
47. Almeida, A.S.; Journel, A.G. Joint simulation of multiple variables with a markov-type coregionalization model. *Math. Geol.* **1994**, *26*, 565–588. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.