minerals

MDPI

Article

Geometallurgical Responses on Lithological Domains Modelled by a Hybrid Domaining Framework

Yerniyaz Abildin [1,2], Chaoshui Xu [1,2], Peter Dowd [1,2,*] and Amir Adeli [1]

[1] PRIF Mining Consortium, Institute for Sustainability, Energy and Resources, University of Adelaide, Adelaide 5005, Australia; yerniyaz.abildin@adelaide.edu.au (Y.A.)
[2] School of Chemical Engineering, University of Adelaide, Adelaide 5005, Australia
* Correspondence: peter.dowd@adelaide.edu.au

Abstract: Identifying mineralization zones is a critical component of quantifying the distribution of target minerals using well-established mineral resource estimation techniques. Domains are used to define these zones and can be modelled using techniques such as manual interpretation, implicit modelling, and advanced geostatistical methods. In practise, domaining is commonly a manual exercise that is labour-intensive and prone to subjective judgement errors, resulting in a largely deterministic output that ignores the significant uncertainty associated with manual domain interpretation and boundary definitions. Addressing these issues requires an objective framework that can automatically define mineral domains and quantify the associated uncertainty. This paper presents a comparative study of PluriGaussian Simulation (PGS) and a Hybrid Domaining Framework (HDF) based on simulated assay grades and XGBoost, a machine-learning classification technique trained on lithological properties. The two domaining approaches are assessed on the basis of the domain boundaries produced using data from an Iron Oxide Copper Gold deposit. The results show that the proposed HDF domaining framework can quantify the uncertainty of domain boundaries and accommodate complex multiclass problems with imbalanced features. Geometallurgical models of the Net Smelter Return and grinding time are used to demonstrate the effectiveness of HDF. In addition, a preprocessing step involving a noise filtering method is used to improve the performance of the ML classification, especially in cases where domain boundaries are difficult to predict due to the similarity in geological characteristics and the inherent noise in the data.

Keywords: domain modelling for resource estimation; geological uncertainty; geostatistical simulation; machine learning; classification; noise filtering; geometallurigical modelling

Citation: Abildin, Y.; Xu, C.; Dowd, P.; Adeli, A. Geometallurgical Responses on Lithological Domains Modelled by a Hybrid Domaining Framework. *Minerals* **2023**, *13*, 918. https://doi.org/10.3390/min13070918

Academic Editor: Simon Dominy

Received: 20 April 2023
Revised: 14 June 2023
Accepted: 15 June 2023
Published: 7 July 2023

1. Introduction

Geometallurgical modelling integrates metallurgical responses into a block model for mine planning. Geometallurigical parameters, such as grinding time and recovery, enable a more exhaustive approach to resource modelling and complement conventional grade and geology-based attributes [1–3]. Introducing metallurgical responses into the modelling process generally provides a more realistic optimisation of the mine's economic objectives with better processing performance assessment and overall project evaluation [4]. Geometallurgical variables provide a basis for improving the net present value, comminution performance, and processing recovery. Thus, early-stage prediction of the variables is critical for effective mine planning and risk assessment.

Each deposit requires a specific approach to geometallurgical modelling based on identifying the geological, geomechanical, mineralogical, and geochemical properties of the ore and connecting them as "input" to the processing "output" [5]. Geostatistical approaches are commonly used to model geometallurgical variables and produce high-resolution maps [6]. However, these variables may have characteristics that make the modelling process tedious. For example, some geometallurgical variables are non-additive,

whereas others have complex multivariate relationships [7–9]. In addition, geometallurgical modelling is often hindered by inappropriate data collection, testing, and analysis and, in some cases, the absence of relevant data. Consequently, there are various challenges in predicting metallurgical responses that can be included in a 3D block model, starting from sample collection to modelling multiple variables with multivariate complexities.

In the geoscience sense, domain modelling (domaining) plays an important role in reliable resource estimation and in establishing a geometallurgical model. In mining applications, domaining methods can be categorised into explicit, implicit, and geostatistical approaches. Briefly, explicit modelling is manual designing, or drawing, in 2D or 3D space, which is usually done by the mine geologist. Implicit modelling was introduced by Cowan et al. [10] and uses radial basis functions to predict domain boundaries. There are various forms of implicit modelling, all of which use prediction techniques such as the potential field method, machine learning algorithms, and combining implicit functions with explicit data [11–13]. There are various geostatistical methods used in the mining industry for building geological, mineralization, and lithological models. The most common approaches are indicator kriging, sequential indicator simulation, and PluriGaussian simulation [14–18]. In addition, geostatistical methods with specific modifications have been developed for domaining purposes. As an example, Emery et al. [19] used a PluriGaussian simulation with spatially varying proportions and tested it on a porphyry copper deposit. For modelling a larger number of geological domains, the hierarchical truncated pluriGaussian simulation is a robust alternative to the original pluriGaussian simulation, which allows the user to control the contact relationships of domains [20,21]. In another example, Séguret [22] developed a "partial grade" estimation method to identify the border effect between domains. There are also hybrid approaches that use geostatistical and machine learning algorithms (e.g., classification and clustering) to predict the boundaries of geological and geometallurgical domains [23–27]. However, most of these domaining methods use one deterministic inter-operation, and thus the uncertainty of the domain boundaries may not be fully quantified [23,28]. Examples of methods and frameworks that provide quantified uncertainty of domain boundaries are hierarchical simulation with signed distance [29] and hybrid frameworks of simulation and machine learning [30,31]. Commonly, only a single deterministic domain model is used for resource estimation during domaining; nevertheless, the quantification of the uncertainty of domain boundaries is essential for mid-term, long-term, and life-of-mine planning [32].

In the work reported here, the aim is to analyse geometallurgical models on the basis of the domains that are produced by PGS and the modified HDF proposed in Abildin et al. [30]. This framework consists of geostatistical simulation and machine learning classification with prior preprocessing. In the first part of the framework, the assay grades are conditionally simulated in a selected region of the deposit. In the second part, the pre-trained classification algorithm predicts the lithological classes of the simulated values in the block model. Preprocessing, or noise filtering, improves the performance of the classifier and, in the case studies with complex data, has been shown to produce reliable results. The Net Smelter Return (NSR) and the particular lithologies are important for comminution and will be used as geometallurgical variables to assess the domaining methods listed above.

The structure of the paper is as follows: Section 2 outlines the methodology proposed for the domaining framework and PGS. In Section 3, we compare and analyse case study material and the domaining procedures for the reference PGS and HDF with different configurations. In Section 4, we discuss the geometallurgical domaining results. Finally, in Section 5, we summarise the conclusions drawn from this study.

2. Methodology

This section presents a comprehensive description of the methods employed in this study. Geostatistical techniques, including PluriGaussian Simulation (PGS) and Turning Bands Simulation (TBSim), were used to model geological categorical and continuous

variables. The factorization technique was used to correlate the continuous variables. Classification and noise filtering were achieved through the implementation of machine learning algorithms. Finally, Section 2.6 outlines the workflow employed for domain modelling and its effects on geometallurgical responses.

2.1. PluriGaussian Simulation

PluriGaussian simulation (PGS) is an extension of Truncated Gaussian Simulation (TGS) and is designed to address a broader range of subsurface models with complex geological contacts between categorical variables such as lithofacies, lithology, and rock types [33]. In particular, it is suited for scenarios with no inherent ordering between categories or in conditions with impermissible contacts. PGS is widely used in oil and gas applications for determining geological domains such as lithofacies [14,34]. PGS has also been applied to modelling lithological units in mining applications [18,19,35,36] and in interface modelling in petroleum applications [14].

The TGS method involves truncating a Gaussian random field using a fixed set of thresholds, with categories of a variable defined as a collection of points whose corresponding Gaussian values fall within the interval defined by the thresholds. In contrast, PGS uses a multivariate truncation rule to truncate multiple Gaussian random fields. This rule partitions a two-dimensional (or higher-dimensional) space into sub-classes that relate the Gaussian values of the random fields to the lithofacies domains. The determination of category locations within a model depends on their contact relationships, which in turn reflect their chronological or topological configurations [25]. A three-domain model with mutually connected categories can be simulated using two Gaussian random fields (a two-dimensional flag) with two threshold values. In the typical practise of lithology or lithofacies modelling, two or three independent Gaussian random fields are used to facilitate the inference of the modelling parameters. The threshold values play a critical role in determining the proportions of each category, where a threshold of zero results in an equal proportion of the two domains, and the Gaussian random field takes on values equally above and below the threshold value. Category proportions based on sample data are used to calculate the threshold values [14,37].

In the context of Gaussian random fields, the covariance functions (or variograms) are associated with the variograms of categorical indicators and describe the spatial correlation properties of the random fields. The categorical data are transformed into indicator values (zero, one, or unknown) based on the Gaussian random fields and their threshold values. The indicator variogram (Equation (1)) of the domain [D] and cross-variograms (Equation (2)) between pairs of categorical domains [$D1$ & $D2$] are then calculated using a non-centred covariance function [38], and the Gaussian variograms are obtained by fitting the indicator variograms through a trial-and-error approach [14,39]:

$$\gamma_D(x,h) = p_D - C_D(h) \tag{1}$$

$$\gamma_{D1D2}(x,h) = -\frac{1}{2}[C_{D1D2}(h) + C_{D2D1}(h)] \tag{2}$$

where h—lag distance, γ—indicator (cross)-variograms, C—indicator (cross)-covariance function. Note that the equations are for domains with constant proportions.

In summary, the remaining steps after inferring the variograms are: first, the categorical indicators in the drillholes are simulated through the Gibbs sampler, where the existing indicator values are converted to Gaussian values. Second, the Gaussian random fields are simulated at the target locations using methods such as TBSim, sequential Gaussian simulation, and the spectral method [40]. Finally, the Gaussian values are truncated on the basis of the pre-defined contact rules to obtain the simulated categories at the target locations (Figure 1).

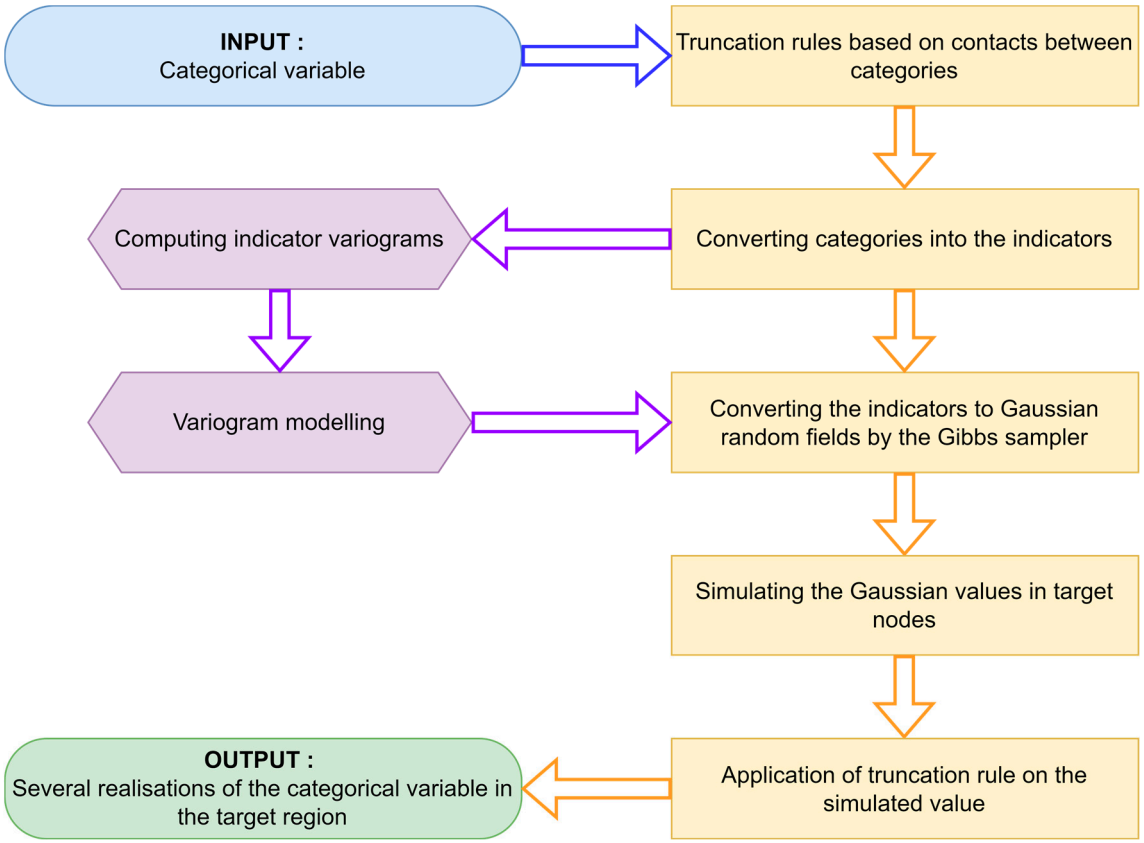

Figure 1. Brief workflow for PGS steps.

2.2. Turning Bands Simulation

Turning Bands Simulation (TBSim) is a technique for simulating a Gaussian random field $Y(x)$ in R^d (where $d \geq 1$) as a weighted sum of one-dimensional (1D) simulation values. The data are first transformed to Gaussian values using the normal score method and projected onto random lines distributed over a unit sphere or circle, depending on the dimensionality of the problem. Next, 1D Gaussian random fields are obtained by multiplying the existing random field with unit vectors [37]. The covariance functions of the 1D Gaussian fields are mapped one by one from the isotopic continuous covariance function in R^d to the continuous covariances in R. The simulations along the 1D vectors are generated using the continuous spectral method, which involves simulating Gaussian random fields using sinusoid functions with random frequency and phases [41]. The frequency of the sinusoid functions is chosen randomly from a distribution derived from the inverse Fourier transform of the covariance function of $Y(x)$ [40]. Once the frequency and phases are determined, a non-conditional simulation of 1D Gaussian random fields can be generated, after which 1D simulations are combined to generate a non-conditional simulation in R^d. The conditioning step is executed by computing the simple kriging of the differences between the data values and the simulated values at the data locations and adding the results to the non-conditionally simulated values. Finally, the conditional simulation results are back-transformed from the normal score scale to the original scale. For more information on TBSim, refer to Emery and Lantuéjoul [37,42].

2.3. Projection Pursuit Multivariate Transform

Projection Pursuit Density Estimation was first introduced by Friedman and Tukey [43] as a method for finding "interesting" projections of high-dimensional data by seeking projections that maximise certain statistical criteria. Barnett et al. [44] developed the Projection Pursuit Multivariate Transform (PPMT) algorithm for multivariate geostatistical modelling, particularly for geostatistical applications in which high-dimensional spatial data can be

challenging to visualise, analyse, and simulate. PPMT has several advantages over other multivariate techniques, such as Principal Component Analysis and Factor Analysis. It can handle multivariate complexities: non-linear and heteroscedastic relationships and specific geological constraints. The PPMT methodology is effective for decorrelating assay grades and can be applied in various geological and mining contexts [30].

The general methodology of PPMT involves three steps: normal score transformation, sphering, and projection pursuit [44]. First, the data matrix is transformed into normal scores using the Gaussian cumulative distribution function. Second, the normal score data is centred, and an orthogonal covariance matrix is computed. Finally, the projection pursuit algorithm is applied to the sphered data to find the optimal projection that maximises the projection index. The iteration process continues until the transformed data are more Gaussian than 95% of the random Gaussian samples, as determined by the convergence threshold.

2.4. Extreme Gradient Boosting

Chen and Guestrin [45] developed a robust machine-learning algorithm entitled Extreme Gradient Boosting (XGBoost) that has gained widespread attention in various fields, including geology. The XGBoost algorithm uses a gradient boosting technique to iteratively add weak decision trees to the model, with each new tree focused on correcting the errors of the previous ones. The algorithm also includes regularisation terms to prevent overfitting and improve the general performance of the model. The XGBoost algorithm has several advantages over other machine learning algorithms, including its scalability, flexibility, and speed.

XGboost has been shown to perform well in various geological classification tasks, such as lithology and mineral identification, by accurately identifying complex patterns in geological data. One of the key applications of XGBoost in geology is lithology identification from drillholes and well logs. Several studies have shown the effectiveness of XGBoost in accurately identifying different lithological formations. For example, Gu et al. [46] used XGBoost to classify lithological facies. In addition to lithology identification, XGBoost has been used in various other geological applications, such as mineral identification and geochemical anomaly detection.

In summary, XGBoost is a robust machine-learning algorithm that has gained popularity in various geological applications. The ability of the algorithm to identify complex patterns in geological data accurately makes it, inter alia, a promising tool for lithology and mineral identification.

2.5. Preprocessing: HybridRepairFilter

The HybridRepairFilter (HRF) algorithm, named by Morales et al. [47] and proposed by Miranda et al. [48], aims to improve the performance of classifiers in the presence of noisy data. The algorithm employs an ensemble of four classifiers, including Support Vector Machines (SVMs), Artificial Neural Networks (ANNs), CART, and k-Nearest Neighbour (KNN) with k values of 1, 3, and 5. The predictions of the classifiers are combined using consensus voting schemes to identify a subset of instances that can be labelled as noise. Depending on the data characteristics, the noisy instances are either removed and/or repaired (i.e., re-labelled) by changing their class to the most voted class in the ensemble. The KNN vote decides whether to remove or repair the instance.

The HybridRepairFilter algorithm iteratively processes the dataset until the accuracy of the ensemble trained with the processed dataset increases. The aim of the algorithm is to find the optimal balance between noise removal and noise repair, resulting in improved classification performance on the original dataset [48].

2.6. Workflow of the Study

This study provides a comparative analysis of a Hybrid Domaining Framework (HDF) previously developed by the authors [30] with specific modifications, together with PGS

as a benchmark for domain modelling. The initial HDF incorporates simulations and classification algorithms with a preprocessing step. It has been used for spatial prediction of multiclass problems, but the preprocessing was performed for a binary problem. In other words, only two classes out of all classes in the predicted domain were preprocessed because those were the problematic classes with high misclassification rates in the case study [30]. In this study, the preprocessing step is extended to the multiclass problem. The same HybridRepairFilter is used for noise filtering but is applied only to a balanced binary problem. Thus, the one-to-one approach is introduced by splitting multiple classes into pairs, and then HybridRepairFilter is used on each pair by considering it as a binary problem and balancing it using the Synthetic Minority Over-Sampling Technique (SMOTE) [49].

In the previous study, only the isotopic dataset was considered in the decorrelation process before grade simulation. As a result, the total number of samples for each variable was decreased for the simulation input. In the work reported here, an imputation technique was used to include all available information, so missing values in a dataset are imputed prior to the decorrelation. Note that the imputed values are removed from the dataset before it is used for subsequent independent grade simulations. In addition, it should also be noted that HDF exclusively uses simulated grade values for each block to predict (i.e., classify) categorical variables, disregarding the categorical values of adjacent blocks. Consequently, the spatial continuity of the predicted block categories is implicitly maintained during the simulation process through the use of soft boundaries for simulating continuous grade variables.

The impact of the above-mentioned changes in the HDF will be compared with the original HDF proposed by the authors and PGS. The results will be assessed in terms of domain boundary changes and their influence on geometallurgical models. Note that an imbalanced input means unequal distributions of class proportions in the training data, and a balanced input means these proportions are equal, or are made equal, by oversampling or undersampling techniques. Overall results are compared using the following methods and settings for both isotopic and heterotopic datasets:

- PluriGaussian simulation—PGS;
- Hybrid domaining framework with imbalanced input for the classifier—HDF imbalanced;
- Hybrid domaining framework with balanced input for the classifier—HDF + SMOTE;
- Hybrid domaining framework with noise filtered and balanced input for the classifier—HDF + HRF + SMOTE.

3. Application to a Case Study

3.1. Case Study Overview

The proposed domaining approach is demonstrated in a case study of an Iron Oxide Copper Gold (IOCG) deposit located in Australia. IOCG deposits are mineral deposits that have become increasingly important due to their high economic value and complexity. South Australia is known to have several world-class IOCG deposits, such as the Olympic Dam, Prominent Hill, Carrapateena, and Hillside deposits [50]. IOCG deposits are characterised by their high concentrations of copper, gold, and other valuable metals, such as silver, uranium, and molybdenum [51].

IOCG deposits are typically formed by hydrothermal fluids sourced from deep within the Earth's crust and deposited in Proterozoic-aged iron oxide-rich sedimentary and volcanic rocks [50]. These fluids deposit metal-rich minerals, such as chalcopyrite, bornite, magnetite, and hematite, in the host rocks, resulting in high-grade ore bodies that can be structurally complex and challenging to mine [51]. Although the complex nature of these deposits presents challenges for exploration and mining, they have been addressed through the development of new techniques and methodologies, such as geometallurgy.

In this case, the geological data contains assay data for numerous elements obtained from drillholes and their corresponding geological features, which have been analysed and interpreted by geologists working at the mine site. For each sample, there are 28 continuous grade variables and one categorical lithology variable. The lithology variable is used to

identify domains for the proposed domaining framework. Despite having over 120 unique lithology categories, most of them can be disregarded because they have an insignificant number of samples or can be merged into more significant categories based on their geological features and statistical characteristics. In this study, four major lithology categories obtained by combining 12 major classes (Table 1) and 20 grade assays (Table 2) were used per the mine geologist's recommendation. The grade variables in Table 2 are considered heterotopic, as assay values for some variables are not available for all samples. The lithology data are imbalanced by major classes such as HMBX, DOLM, and OTHR and minor classes such as HQBX (Table 1). The grade values are based on a composite length of two metres, and the lithology for each sample is selected from the prevailing category within the composite interval. As class OTHR contains several lithological categories, it has broader statistical characteristics, which may overlap with the other three categories. In order to show essential statistical differences between the classes with respect to assay grades, Figure 2 illustrates examples of variables with distinct statistical characteristics. Consequently, continuous and categorical variables are suitable for the classification algorithm of the proposed domaining method, which relies on inter-relationships among the variables.

Table 1. Brief description and statistical characteristics of the lithologies.

No.	Category	Name	Heterotopic Data		Isotopic Data	
			Number of Samples	Proportion (%)	Number of Samples	Proportion (%)
1	HMBX	Hematite Breccia	37,755	28.74	25,966	28.01
2	HQBX	Hematite Quartz Breccia	3510	2.67	1182	1.27
3	DOLM	Dolomite	31,506	23.99	25,809	27.84
4	OTHR	All other classes	58,574	44.60	39,749	42.88

Table 2. Brief description and statistical characteristics of the grade assays (pct—percentage, ppm—parts per million).

No.	Abbreviation	Description	Number of Samples	Min.	Max.	Mean	Variance
1	Al_pct	Aluminium	125,779	0.0	14.8	3.74	7.34
2	Au_ppm	Gold	131,343	0.0	137.6	0.34	1.58
3	Ba_pct	Barium	131,261	0.0	13.1	0.31	0.38
4	Ca_pct	Calcium	128,036	0.0	27.2	4.12	30.6
5	Ce_ppm	Cerium	131,310	2.0	1.7×10^4	817.80	8.2×10^5
6	Cr_ppm	Chromium	131,261	1.0	1.6×10^3	62.79	6.4×10^3
7	Cu_pct	Copper	131,345	0.0	20.6	0.4	0.89
8	F_pct	Fluorine	102,786	0.0	13.2	0.27	0.15
9	Fe_pct	Iron	131,311	0.4	69.5	19.87	165.4
10	K_ppm	Potassium	131,064	25	8.6×10^4	1.8×10^4	2.0×10^8
11	La_ppm	Lanthanum	131,258	2.0	1.2×10^4	5.5×10^2	3.9×10^5
12	Mg_ppm	Magnesium	131,261	25	1.7×10^5	2.8×10^4	9.5×10^8
13	Mn_ppm	Manganese	128,245	10	3.3×10^4	2.3×10^3	8.8×10^6
14	Na_ppm	Sodium	131,261	20	6.0×10^4	1.5×10^3	1.8×10^7
15	P_ppm	Phosphorus	131,261	2.5	5.5×10^4	1.3×10^3	1.2×10^6
16	S_pct	Sulphur	127,799	0.0	15.3	0.37	0.30
17	Si_pct	Silicon	125,653	0.1	44.9	18.56	64.4
18	Ti_ppm	Titanium	131,261	5.0	2.9×10^4	3.7×10^3	1.4×10^7
19	Zn_ppm	Zinc	131,345	1.0	1.0×10^4	35.4	4.7×10^3
20	Zr_ppm	Zirconium	93,100	2.5	862.5	125.3	1.0×10^4

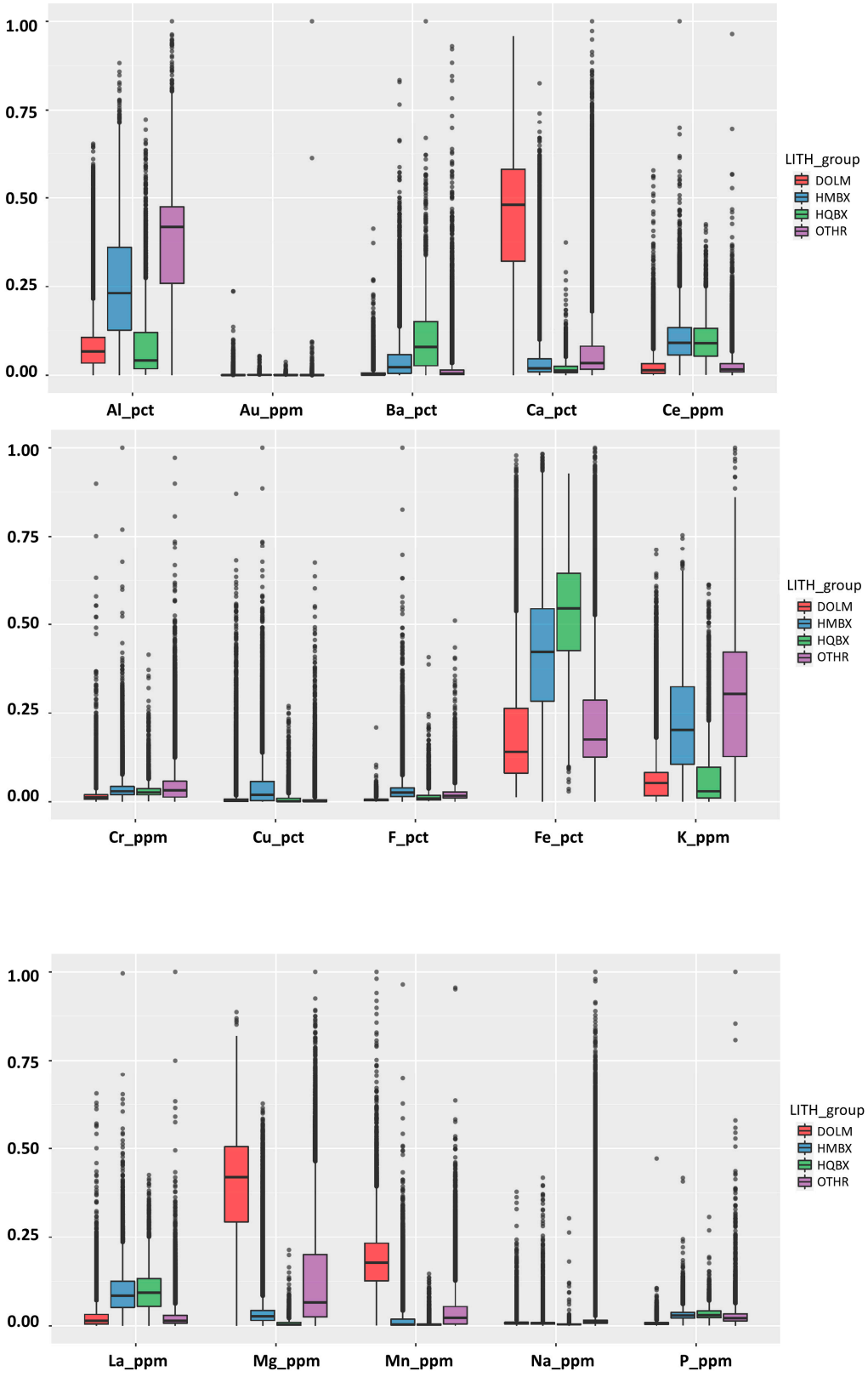

Figure 2. *Cont.*

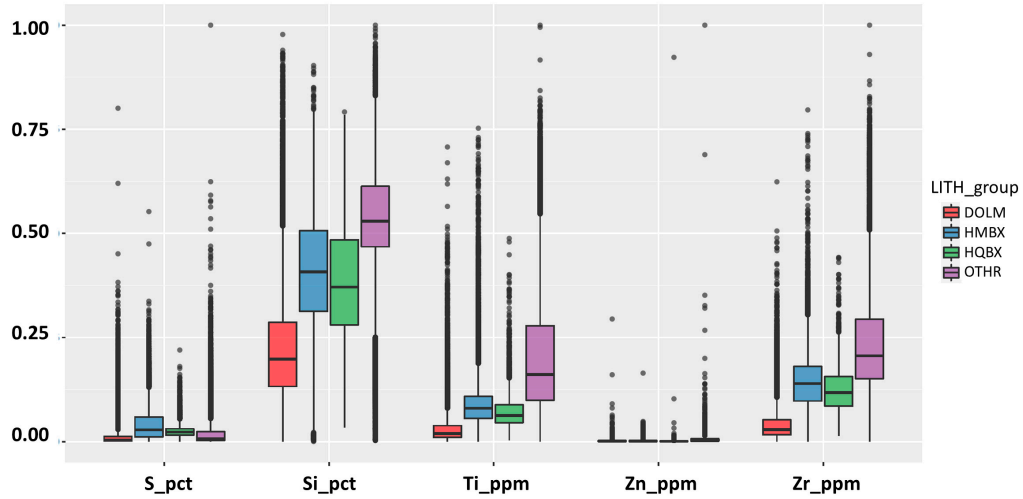

Figure 2. Standardised boxplots for the variables grouped by lithological categories.

3.2. PluriGaussian Simulation

The initial stage of PGS requires the definition of the truncation rule and the specification of the threshold values. In this approach, the domains are characterised by truncating a set of Gaussian random fields. The determination of the number of Gaussian random fields, as well as the method of truncation, depends on the contact relationship. In our case, the mine geologist recommended the truncation rule shown in Figure 3.

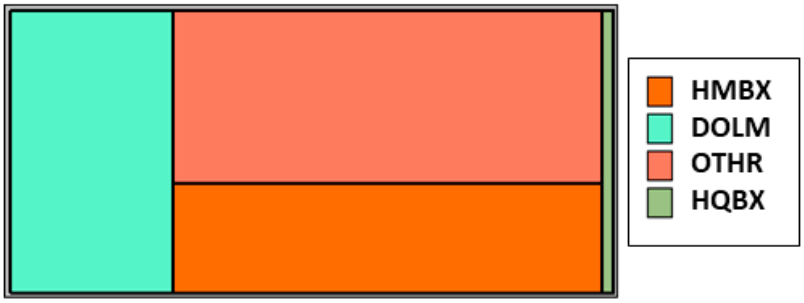

Figure 3. Truncation rule for the lithologies while maintaining the relative proportions.

After determining the threshold values for isotopic and heterotopic data, the indicator data were generated. To determine the variogram models for the Gaussian random fields, experimental variograms were computed for the indicator data in both horizontal and vertical directions. Based on the experimental indicator variograms, the variogram models for the Gaussian random fields were inferred and fitted. It is important to note that the cross-variograms of the Gaussian random fields were assumed to be zero. The fitted Gaussian variograms are shown in Figure 4. The modelled variograms comprised two nested structures (Table 3). Once the variogram models were inferred, the next step involved converting the indicators into Gaussian data using the Gibbs sampler. It involves iteratively assigning categories to locations on a simulation grid based on conditional distributions defined by input data. By considering the spatial characteristics of each category and their interrelationships, the Gibbs sampler produces realistic realisations that capture the complexity and uncertainty of geological systems. Subsequently, 100 realisations were produced and truncated in accordance with the defined truncation rules mentioned above.

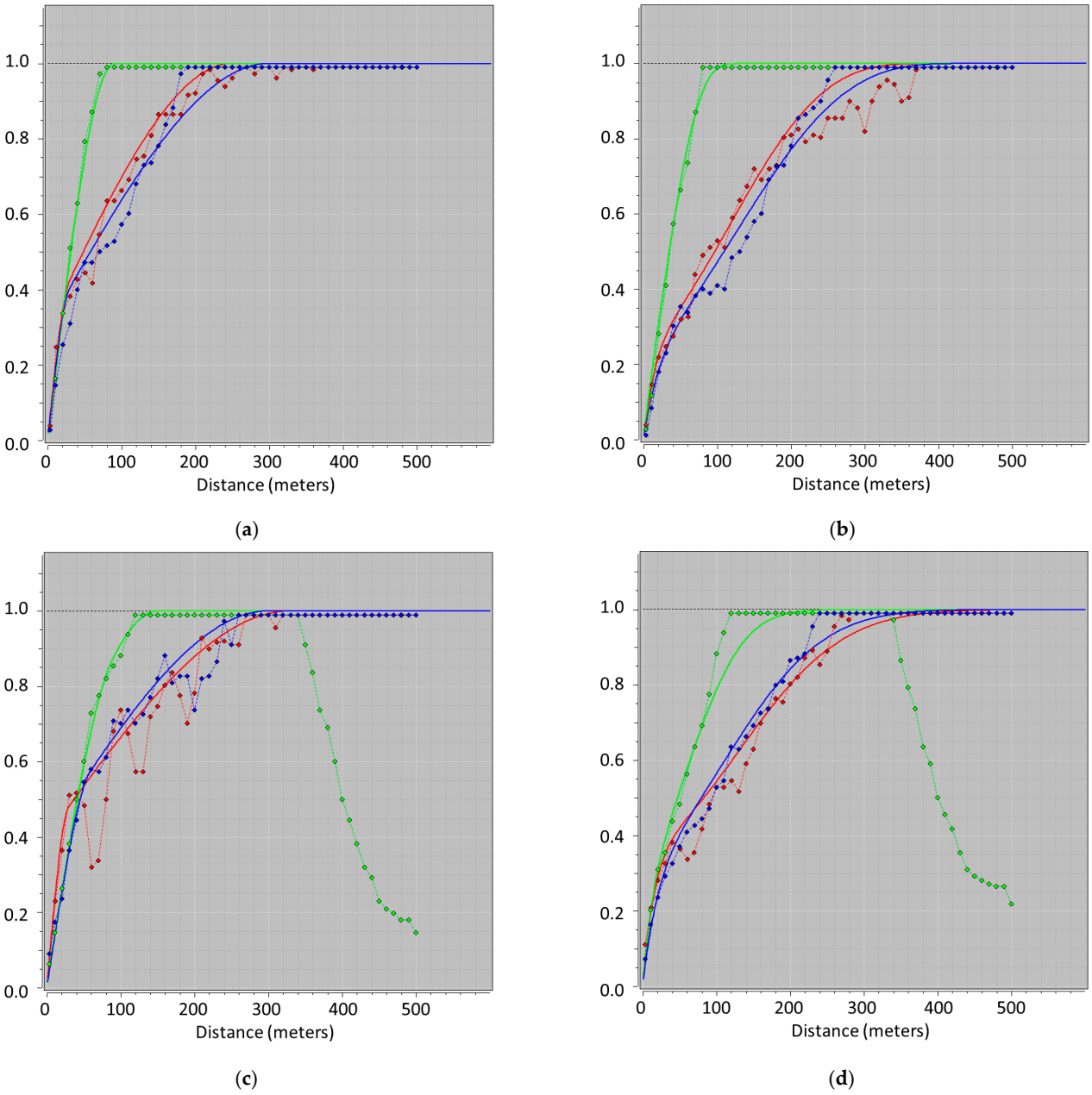

Figure 4. Gaussian variogram models: the experimental variograms are indicated by points and the fitted models ((**a**,**c**)—Isotopic; (**b**,**d**)—Heterotopic) are the solid lines in the East (red), the North (green) and the vertical (blue) directions. (**a**) Y(Isotopic—1). (**b**) Y(Heterotopic—1). (**c**) Y(Isotopic—2). (**d**) Y(Heterotopic—2).

Table 3. Parameters of the fitted Gaussian variograms models.

No.	Gaussian Random Field	Nugget	Structure 1				Structure 2			
			c	*a (EW), m*	*a (NS), m*	*a (vert), m*	*c*	*a (EW), m*	*a (NS), m*	*a (vert), m*
a	γ(Isotopic − 1)	0	0.30 Sph	30	90	30	0.70 Sph	250	90	300
c	γ(Isotopic − 2)	0	0.40 Sph	30	90	60	0.60 Sph	330	150	300
b	γ(Heterotopic − 1)	0	0.30 Gauss	50	50	70	0.70 Cub	400	130	450
d	γ(Heterotopic − 2)	0	0.40 Exp	50	50	50	0.60 Exp	330	170	300

3.3. Hybrid Domaining Framework

3.3.1. Modelling and Simulation of Continuous Grade Variables

The PPMT transformation used for the decorrelation process reduces the initial heterotopic dataset of 131 k samples to an isotopic dataset of 93 k samples. This transformation leads to the elimination of the heterotopic portion of the whole dataset, which means disregarding some useful information. However, data imputation can mitigate the issue of unequal sampling, so the complete set of original data can be used for the simulation [44]. In this study, we used the MissForest machine learning-based imputation technique to infer missing values [52]. As a result, by retaining the original number of samples, the heterotopic dataset becomes isotopic. Both isotopic datasets with 131 k and 93 k samples, respectively, are evaluated in terms of domain boundary changes and HDF performance. However, before TBSim is applied, the imputed values are removed because they might be a source of potential bias, thus introducing artificial variability and impacting model performance. In other words, the imputation impacts only the PPMT forward and back transformations and indirectly impacts the simulation results because previously unused information is retained in the domaining process. Therefore, the excluded samples from the isotopic dataset are fully employed in the independent simulation of the heterotopic dataset. Consequently, the uncertainty of the locations with the retained samples is reduced for the corresponding variables.

The cross-correlated variables are transformed into independent Gaussian variables using PPMT. The histograms of six out of the twenty grade variables are presented to provide a better visual understanding of their original distribution (Figure 5). The variables are not segregated by specific criteria such as mineralogy or rock types. Consequently, some grade variables exhibit multi-modal distributions due to the presence of different minerals for the same element. The histograms in Figure 5 show the results of the PPMT transformation examples, which are similar for the remaining transformed variables. Figure 6 displays the bi-variate scatter plots of the original and transformed variables for pairs of grade variables with the most interesting inter-relationships. As can be seen, the example pairs are decorrelated and free from multivariate correlations. This is also the case for the other variables. Consequently, the subsequent steps, including variogram modelling and independent simulations, can be taken.

As previously stated, the simulation step regards the entire region as a stationary zone for variography and assumes the presence of gradual boundaries between lithological domains. It should be noted that the proposed domaining method partitions the region into these domains during the classification step. Following the forward transformation, variogram analyses were conducted on the transformed data, and Tables 4 and 5 display the resulting fitted variogram models for the isotopic and heterotopic data, respectively. In addition, Figure 7 provides a graphical representation of the variogram models for the transformed Fe variable.

Table 4. Fitted variogram models for the grade assay variables for the isotopic subset (pct—percentage, ppm—parts per million).

No.	Grade Variable	Nugget	Structure 1				Structure 2			
			c	*a (EW), m*	*a (NS), m*	*a (vert), m*	*c*	*a (EW), m*	*a (NS), m*	*a (vert), m*
1	$\gamma(Al_{PPMT\_Gauss})$	0.40	0.45	25	17	24	0.15	500	324	500
2	$\gamma(Au_{PPMT\_Gauss})$	0.45	0.48	34	33	41	0.07	274	500	176
3	$\gamma(Ba_{PPMT\_Gauss})$	0.41	0.32	21	37	39	0.27	130	67	130
4	$\gamma(Ca_{PPMT\_Gauss})$	0.32	0.38	22	31	43	0.30	500	207	222
5	$\gamma(Ce_{PPMT\_Gauss})$	0.34	0.40	20	17	26	0.26	203	92	188
6	$\gamma(Cr_{PPMT\_Gauss})$	0.40	0.44	20	18	26	0.16	289	120	380
7	$\gamma(Cu_{PPMT\_Gauss})$	0.28	0.49	32	42	51	0.23	140	50	90
8	$\gamma(F_{PPMT\_Gauss})$	0.24	0.42	19	19	23	0.34	180	112	184

Table 4. *Cont.*

No.	Grade Variable	Nugget	Structure 1				Structure 2			
			c	*a (EW), m*	*a (NS), m*	*a (vert), m*	*c*	*a (EW), m*	*a (NS), m*	*a (vert), m*
9	$\gamma(Fe_{PPMT\_Gauss})$	0.36	0.40	25	30	30	0.24	50	59	140
10	$\gamma(K_{PPMT\_Gauss})$	0.25	0.51	18	14	18	0.24	107	56	184
11	$\gamma(La_{PPMT\_Gauss})$	0.43	0.39	20	21	28	0.18	351	303	500
12	$\gamma(Mg_{PPMT\_Gauss})$	0.35	0.36	23	28	33	0.29	426	284	226
13	$\gamma(Mn_{PPMT\_Gauss})$	0.21	0.41	18	20	28	0.38	165	99	153
14	$\gamma(Na_{PPMT\_Gauss})$	0.28	0.37	10	20	19	0.35	117	81	111
15	$\gamma(P_{PPMT\_Gauss})$	0.35	0.41	27	19	24	0.24	222	87	500
16	$\gamma(S_{PPMT\_Gauss})$	0.27	0.23	19	36	29	0.50	234	126	249
17	$\gamma(Si_{PPMT\_Gauss})$	0.30	0.39	15	11	18	0.31	100	73	120
18	$\gamma(Ti_{PPMT\_Gauss})$	0.25	0.47	18	15	19	0.28	53	48	94
19	$\gamma(Zn_{PPMT\_Gauss})$	0.26	0.43	17	12	17	0.31	105	61	128
20	$\gamma(Zr_{PPMT\_Gauss})$	0.30	0.43	18	14	17	0.27	87	66	102

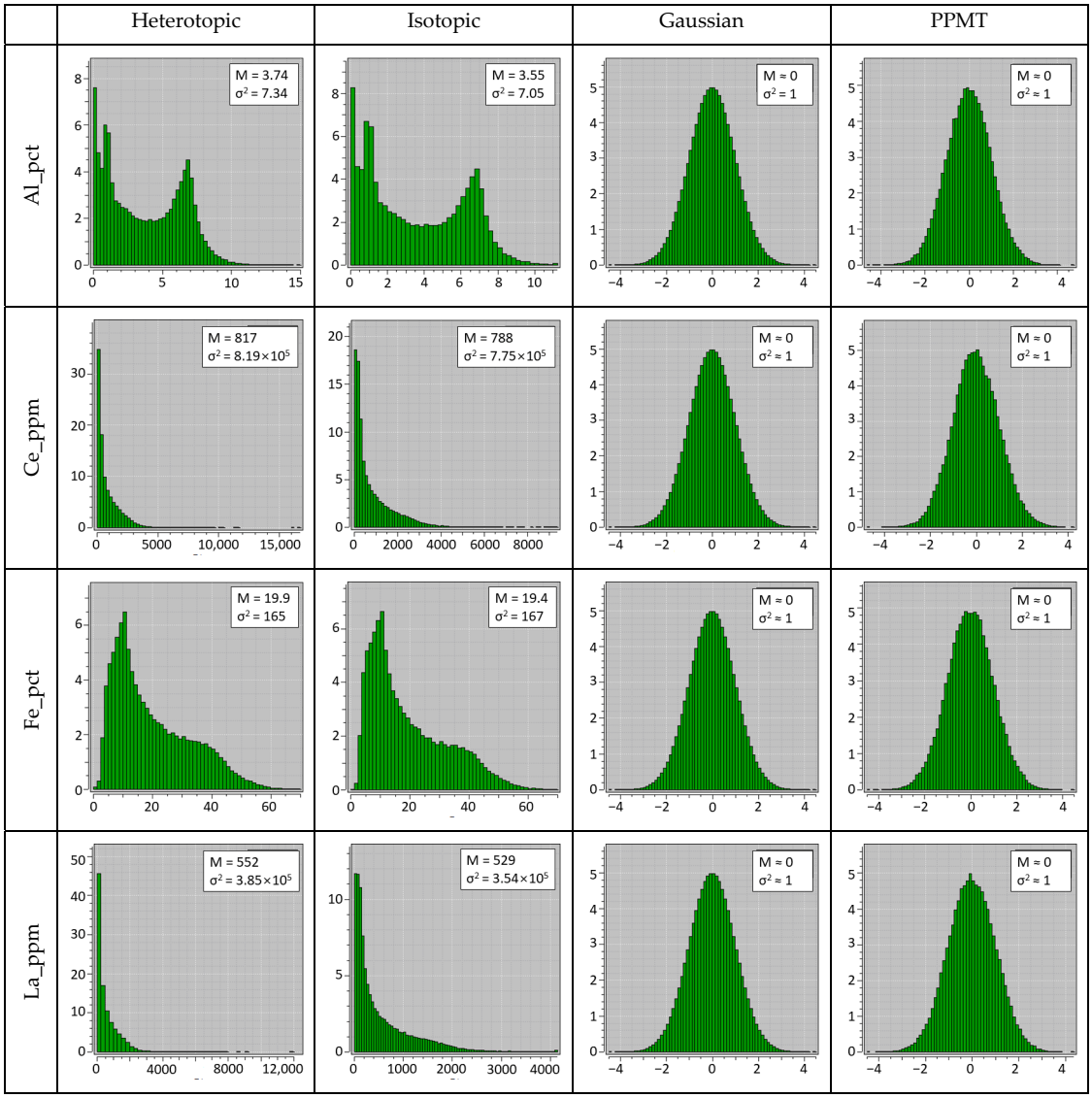

Figure 5. *Cont.*

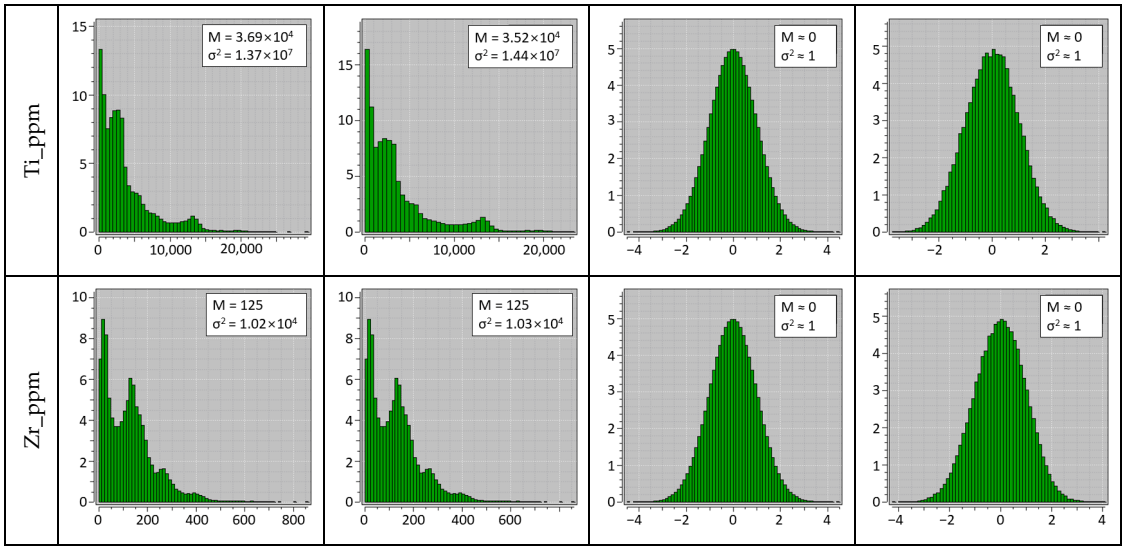

Figure 5. Histograms examples of six grade assay variables: original heterotopic, isotopic, standard Gaussian and PPMT transformed factors based on the isotopic subset (pct—percentage, ppm—parts per million).

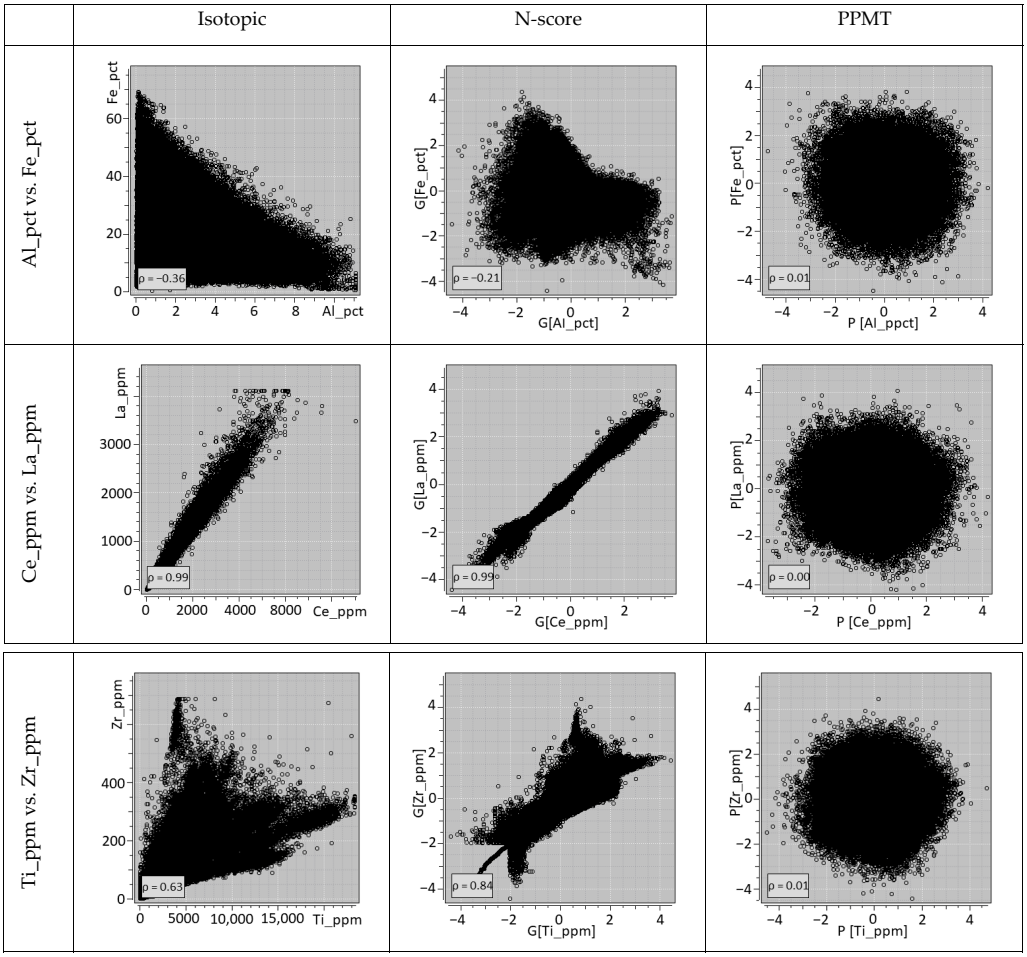

Figure 6. Bivariate scatterplots of Al vs. Fe, Ce vs. La, and Ti vs. Zr of original (**left**), Gaussian (**centre**) and PPMT transformation (**right**) (G[]—Gaussian, P[]—PPMT factor, pct—percentage, ppm—parts per million).

Table 5. Fitted variogram models for the grade assay variables for the heterotopic data (pct—percentage, ppm—parts per million).

No.	Grade Variable	Nugget	Structure 1				Structure 2			
			c	*a (EW), m*	*a (NS), m*	*a (vert), m*	*c*	*a (EW), m*	*a (NS), m*	*a (vert), m*
1	$\gamma(Al_{PPMT\_Gauss})$	0.35	0.46	20	16	23	0.19	344	196	500
2	$\gamma(Au_{PPMT\_Gauss})$	0.41	0.45	32	36	43	0.14	500	500	303
3	$\gamma(Ba_{PPMT\_Gauss})$	0.38	0.32	22	35	35	0.30	163	87	151
4	$\gamma(Ca_{PPMT\_Gauss})$	0.33	0.34	21	26	32	0.33	500	199	310
5	$\gamma(Ce_{PPMT\_Gauss})$	0.33	0.41	24	16	27	0.26	224	101	217
6	$\gamma(Cr_{PPMT\_Gauss})$	0.34	0.46	19	18	26	0.20	358	167	417
7	$\gamma(Cu_{PPMT\_Gauss})$	0.28	0.58	44	43	72	0.14	500	324	167
8	$\gamma(F_{PPMT\_Gauss})$	0.24	0.43	18	15	19	0.33	180	111	210
9	$\gamma(Fe_{PPMT\_Gauss})$	0.34	0.45	30	28	40	0.21	50	65	141
10	$\gamma(K_{PPMT\_Gauss})$	0.25	0.50	19	14	19	0.25	163	130	283
11	$\gamma(La_{PPMT\_Gauss})$	0.39	0.41	19	19	27	0.20	333	245	500
12	$\gamma(Mg_{PPMT\_Gauss})$	0.31	0.35	22	28	31	0.34	480	285	311
13	$\gamma(Mn_{PPMT\_Gauss})$	0.21	0.40	19	20	29	0.39	188	120	203
14	$\gamma(Na_{PPMT\_Gauss})$	0.26	0.43	10	18	22	0.31	117	87	116
15	$\gamma(P_{PPMT\_Gauss})$	0.33	0.43	26	17	24	0.24	375	238	500
16	$\gamma(S_{PPMT\_Gauss})$	0.27	0.23	20	30	27	0.5	220	120	241
17	$\gamma(Si_{PPMT\_Gauss})$	0.28	0.40	15	10	18	0.32	141	73	146
18	$\gamma(Ti_{PPMT\_Gauss})$	0.29	0.47	19	19	25	0.24	112	67	150
19	$\gamma(Zn_{PPMT\_Gauss})$	0.27	0.40	18	16	18	0.33	170	80	213
20	$\gamma(Zr_{PPMT\_Gauss})$	0.33	0.67	31	18	31				

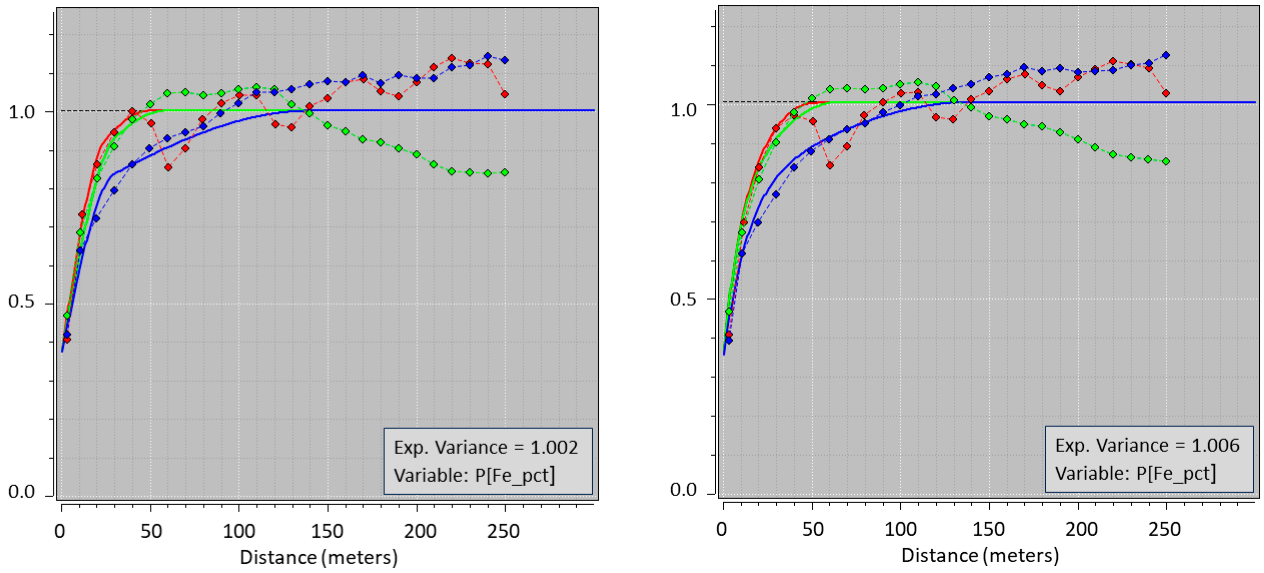

Figure 7. Variogram models of an example variable, Fe_pct(PPMT), for both the isotopic (**left**) and heterotopic (**right**) cases. The experimental variograms are indicated by points and the fitted models by a solid line along the East (red), the North (green) and Vertical (blue) directions (pct—percentage, ppm—parts per million).

The simulations were generated on a regular grid consisting of 80 × 50 × 65 nodes with a 10-metre grid spacing in all directions. Hereafter, the simulated values of the "blocks" in this case are actually values at the corresponding simulation grid nodes. TBSim was used together with the variogram models in Tables 4 and 5 to generate 100 realisations for each variable and subset. Once completed, the simulated values were converted back to the original data space. Figure 8 displays the E-type maps, which indicate the average values of 100 realisations for two example variables, Al and Fe, for both the isotopic and heterotopic cases.

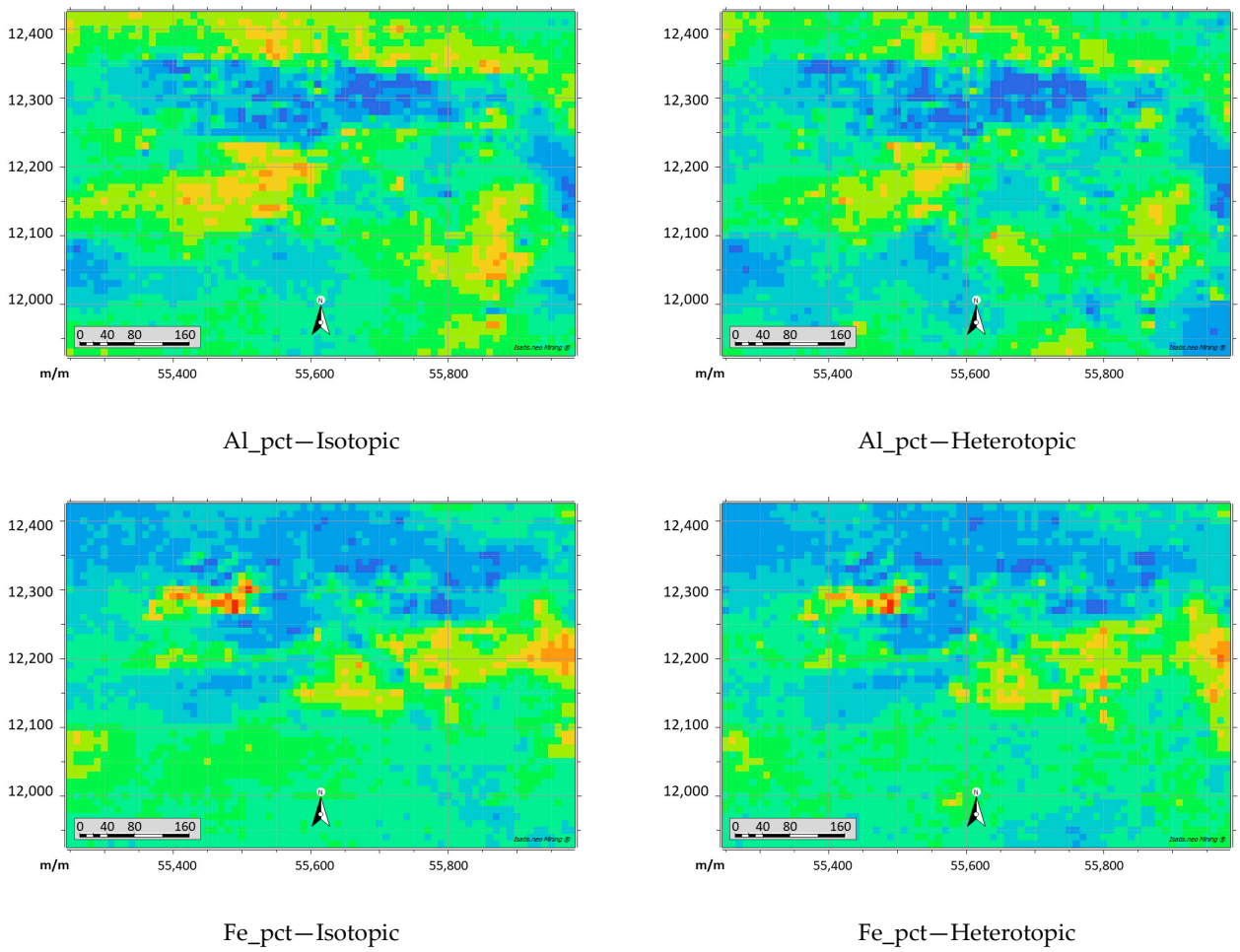

Figure 8. Top view of the cross-section of the average value of 100 realisations for isotopic and heterotopic cases of Al and Fe at the elevation Z = 9715 m.

3.3.2. Validation of Simulation Results

The statistical characteristics of the variables, such as global mean and variance, were analysed to assess the accuracy of the simulated grade values. Mean and variance values for each realisation were obtained, and their average values over 100 realisations were compared with the original values, as presented in Table 6. Validation of the local statistical parameters is another way of checking the quality of the simulation results and assessing the reproduction of spatial continuity. The direct variograms of the 100 simulated realisations should be compared with the fitted variogram model [53]. As an illustration, the direct variograms of the Fe variable obtained from TBSim together with the associated models used are shown in Figure 9a, indicating that the spatial continuity is reasonably well reproduced [54]. In addition, Figure 9b depicts swath plots of the grade distributions derived from several directions through the deposit, including East, North, and Vertical directions for the block model and composites. It should be noted that the average values at both ends of the curves are based on a small number of data points (e.g., Figure 9b: swath

plot—vertical). A comparison of the simulated values and the original data shows mostly similar results, with some slight deviation observed in the swath plots, which could be due to the strongly skewed distributions of the example variable.

Table 6. Global statistical characteristics of the original data and simulation results (pct—percentage, ppm—parts per million).

	Mean				Variance			
Grade Assays	**Isotopic**	**Mean of 100 Realisations**	**Heterotopic**	**Mean of 100 Realisations**	**Isotopic**	**Mean of 100 Realisations**	**Heterotopic**	**Mean of 100 Realisations**
Al_pct	3.55	3.72	3.74	3.76	7.05	5.41	7.34	5.50
Au_ppm	0.33	0.39	0.34	0.40	0.98	1.61	6.30	11.90
Ba_pct	0.27	0.54	0.31	0.48	0.31	1.01	0.38	0.72
Ca_pct	4.82	3.26	4.12	2.86	34.39	16.88	30.63	13.74
Ce_ppm	788.10	798.69	817.76	741.80	7.8×10^5	6.9×10^5	8.2×10^5	7.0×10^5
Cr_ppm	59.03	81.72	62.79	66.59	5.3×10^3	1.1×10^4	6.4×10^3	6.9×10^3
Cu_pct	0.39	0.33	0.40	0.29	0.84	0.59	0.89	0.51
F_pct	0.26	0.30	0.27	0.27	0.13	0.24	0.15	0.16
Fe_pct	19.39	21.30	19.87	21.09	167.60	149.00	165.40	131.30
K_ppm	1.7×10^4	1.7×10^4	1.8×10^4	1.7×10^4	1.9×10^8	1.4×10^8	2.0×10^8	1.4×10^8
La_ppm	5.3×10^2	5.3×10^2	5.5×10^2	5.1×10^2	3.5×10^5	2.9×10^5	3.9×10^5	3.8×10^5
Mg_ppm	3.2×10^4	2.1×10^4	2.8×10^4	1.9×10^4	1.0×10^9	4.8×10^8	9.5×10^8	4.4×10^8
Mn_ppm	2.7×10^3	1.9×10^3	2.3×10^3	1.7×10^3	9.8×10^6	5.3×10^6	8.8×10^6	4.8×10^6
Na_ppm	1.5×10^3	1.9×10^3	1.5×10^3	2.2×10^3	1.5×10^7	2.4×10^7	1.8×10^7	3.2×10^7
P_ppm	1.2×10^3	1.5×10^3	1.3×10^3	1.6×10^3	9.9×10^5	2.4×10^6	1.2×10^6	3.5×10^6
S_pct	0.36	0.42	0.37	0.38	0.27	0.35	0.30	0.40
Si_pct	17.79	20.01	18.56	20.41	66.40	47.19	64.28	42.17
Ti_ppm	3.5×10^3	4.3×10^3	3.7×10^3	4.5×10^3	1.4×10^7	1.6×10^7	1.4×10^7	1.6×10^7
Zn_ppm	35.30	38.94	35.42	40.91	2.6×10^3	5.0×10^3	4.7×10^3	4.6×10^3
Zr_ppm	125.07	151.61	125.34	155.23	1.0×10^4	1.2×10^4	1.0×10^4	1.0×10^4

Bivariate scatterplots and correlation coefficients were used to validate the reproduction of the correlations between grade variables. Figure 10 illustrates the bivariate distributions between three example pairs of variables based on their original and simulated values. The overall correlation structures and bivariate distributions were reasonably well reproduced, with the heat maps of the bivariate scatterplots of simulated values being similar to those of the original data. However, certain simulated values (e.g., Ti vs. Zr) did not demonstrate perfect agreement, which may be attributed to the simulation and back-transformation encountering difficulties in managing some multivariate complexities such as heteroscedasticity and other geological constraints in multivariate geostatistical modelling, including unequal sampling, data quality, and scale. It is worth noting that Figure 10 only represents the bivariate scatterplots for one realisation, and all other variables were subjected to the same process of examination to ensure the accuracy of the simulation within the proposed domaining framework.

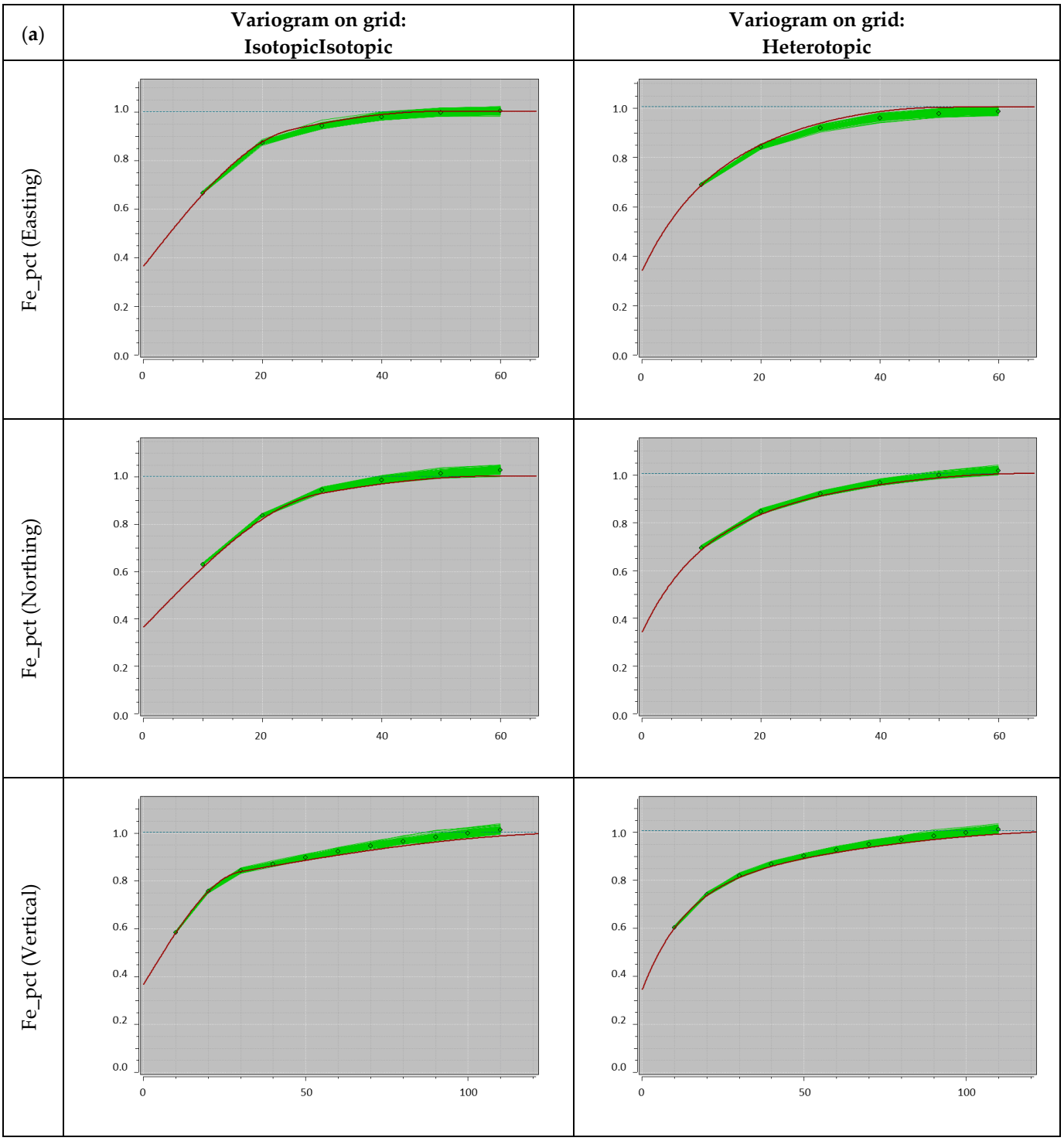

Figure 9. *Cont.*

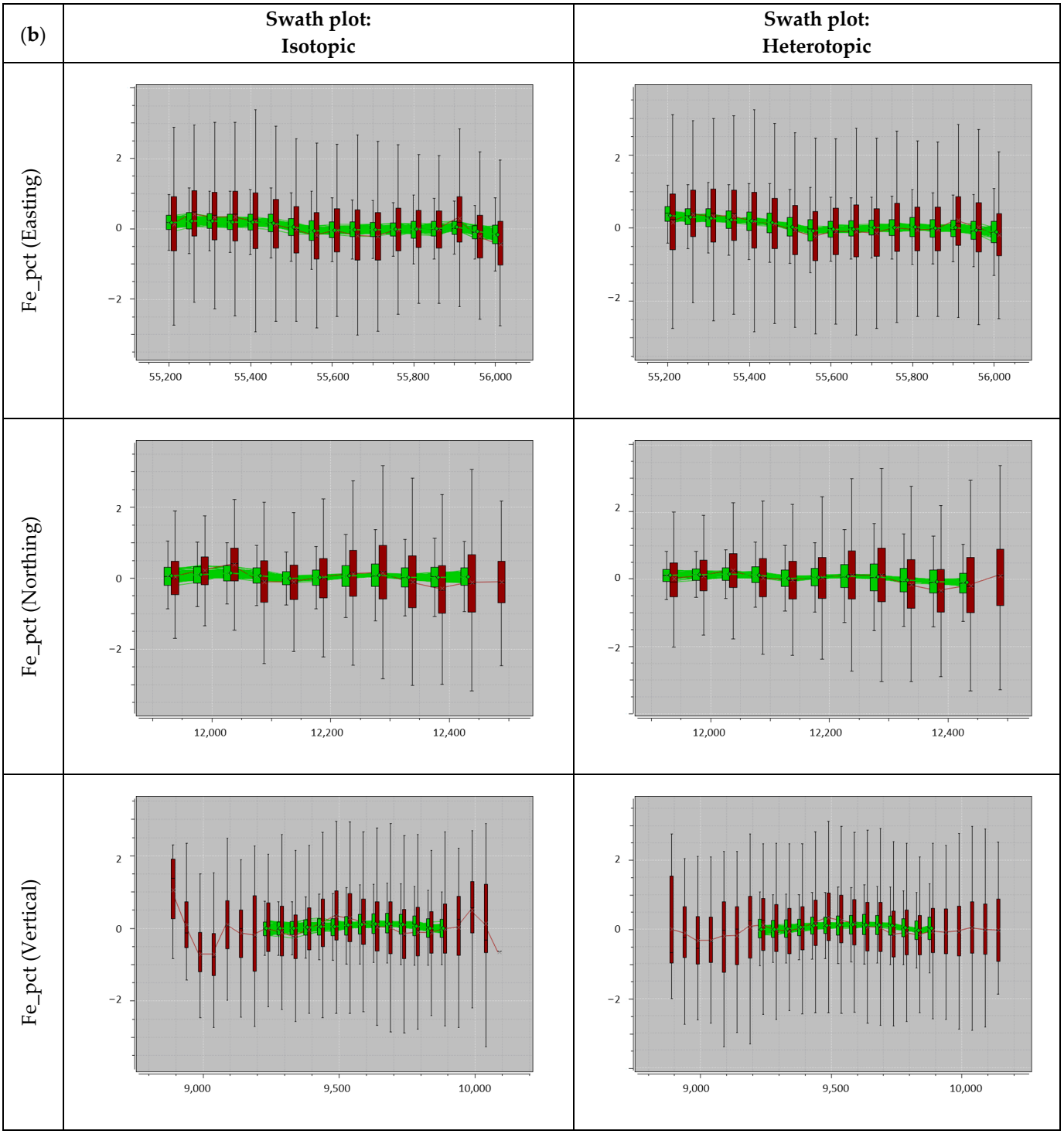

Figure 9. Cross-validation of the original values of Al and Fe: (**a**) variogram of simulated values on the grid (red line—fitted variogram model; green lines and dots—variograms and their mean on the grid) and (**b**) swath plots for (**left**) isotopic and (**right**) heterotopic data in the East, North and Vertical directions (red—original values; green—simulated realisations).

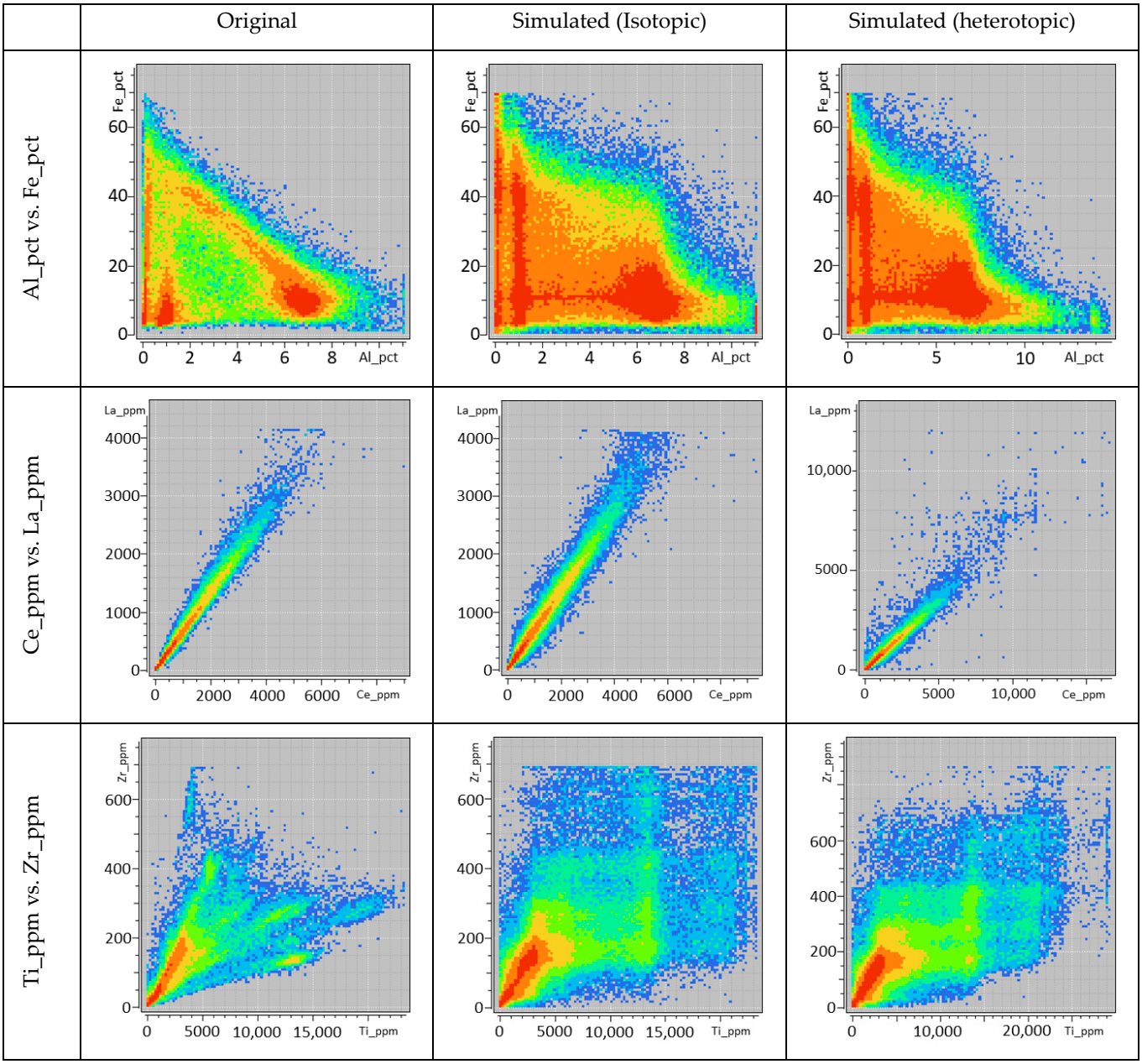

Figure 10. Bivariate scatterplots of Al vs. Fe, Ce vs. La, and Ti vs. Zr for the original (**left**) values, and back-transformed simulated values on the isotopic subset (**centre**) and heterotopic subset (**right**) of Realisation 1 (pct—percentage, ppm—parts per million).

3.3.3. Preprocessing of Noisy and Misinterpreted Lithology Classes in Samples

The uncertainty associated with the domain boundaries is assessed by the simulation part of the proposed framework, which generates 100 realisations for the assay grades. However, the classification algorithms used in the case study struggle to accurately predict certain lithological classes, such as HQBX and OTHR. Based on discussions with on-site geologists, this difficulty is attributed to the complex geological characteristics and potential errors in the geological core logging process. The extensive and diverse geochemical properties of the OTHR class may correspond to several different combined categories, resulting in incorrect classification of some observations in the dataset. Consequently, the overall rate of misclassification increases. Moreover, distinguishing between HMBX and HQBX poses a challenge, as one can easily be misinterpreted as the other. In fact, a considerable number of HQBX samples are misclassified as HMBX, and this misclassification was identified when

examining the overlapping statistical characteristics between the two lithologies [30]. Most of these misclassifications occur in a specific zone in which the block domain model shows HQBX as the second-most likely category after HMBX. However, the expert knowledge of site geologists suggests that the indicated area should contain a significant number of HQBX samples.

Standard machine learning techniques were used to address the problem of incorrect data input by applying noise filtering techniques based on the similarity and dissimilarity of the selected features. In this case study, HRF was employed as an additional step prior to domain classification. As this noise filtering method works only for binary problems, a pairwise approach was used for the multiclass problem. In other words, all possible combinations of pairs among HMBX, HQBX, DOLM, and OTHR classes were processed for the noise filtering. In addition, SMOTE was used to balance each pair to minimise the influence of major classes during the training process. If the balancing is not implemented for highly imbalanced cases, the noise filtering step will be biased towards classes with the highest number of samples [30]. Table 7 provides the statistics of the re-labelled and removed samples in accordance with the corresponding pairs, and Table 8 demonstrates the overall changes of the classes within the isotopic data fed to the XGBoost classifier.

Table 7. The results after applying the noise filtering step for the HMBX, HQBX, DOLM and OTHR classes.

	HQBX		**HMBX**		**DOLM**		**OTHR**	
	Removed	**Re-Labelled**	**Removed**	**Re-Labelled**	**Removed**	**Re-Labelled**	**Removed**	**Re-Labelled**
HQBX	-	-	2029	1953	701	300	1585	865
HMBX	128	0	-	-	839	447	2308	2035
DOLM	30	0	903	328	-	-	3011	1833
OTHR	56	3	1637	705	1343	1775	-	-
Total	214	3	4569	2986	2883	2522	6904	4733

Table 8. Quantitative difference of the lithological classes before and after noise filtering.

	Original		**HybridRepairFilter**	
	Number of Samples	**Proportion (%)**	**Number of Samples**	**Proportion (%)**
HQBX	1182	1.27	4083	5.19
HMBX	25,966	28.01	20,893	26.56
DOLM	25,809	27.84	22,565	28.68
OTHR	39,749	42.88	31,137	39.58

3.3.4. Domain (Lithology) Classification and Uncertainty Quantification

Mine geologists defined the grade domains at the case study mine site based on copper and gold grade values, and the geometries of domain boundaries were manually created. However, due to the subjectivity and uncertainty involved in this approach, a semi-automatic lithological domaining method is of significant interest to the mining operation. Despite the challenges associated with the manual creation of important domains (i.e., HQBX, HMBX, and DOLM), they provide crucial geological information for more precise resource estimation. As a data-driven and semi-automatic approach, the proposed HDF, especially the classification part of the framework, is examined for domain modelling purposes. Thus, XGBoost is trained with several configurations for a better understanding of the classifier's performance and its impact on the final domain model.

The model training and validation process involves several steps, including splitting the dataset into 70/30 training and testing subsets through stratified resampling followed by stratified 3-fold cross-validation, iterative training of the classifiers with different hyperparameters, and testing the classifiers for misclassification rates [30]. The isotopic subset of the original dataset is examined here in terms of classifier performances with the noise filtering method, HRF, and balancing algorithm, SMOTE. Table 9 shows the balanced accuracy obtained during the XGBoost training process. The original imbalanced case is the isotopic data without any changes, whereas the original balanced case involves the application of SMOTE prior to the training process in order to reduce bias towards the minor class, HQBX. The noise-filtered balanced case is based on the result obtained in the previous sub-section and balanced by SMOTE before the training process.

Table 9. The balanced accuracy for the multiclass problem for the original imbalanced, balanced and preprocessed subsets.

	Original Imbalanced	**Original Balanced**	**Noise Filtered Balanced**
Class	Balanced Accuracy	Balanced Accuracy	Balanced Accuracy
HQBX	0.707	0.752	0.866
HMBX	0.863	0.881	0.901
DOLM	0.931	0.937	0.946
OTHR	0.895	0.896	0.913
Average	***0.849***	***0.867***	***0.907***

In the proposed domaining framework, the outputs of the simulations and the classification model are merged in the final stage. The XGBoost classifier, which has been trained for different cases, is used to forecast the lithology domain at the point level using the simulated values of 20 grade variables, resulting in 100 domain realisations. The most probable outputs for each case are depicted in Figure 11 as examples. It is important to note that XGBoost uses 20 simulated grade values exclusively to predict block lithology without explicitly considering the lithology of neighbouring blocks. In other words, XGBoost does not explicitly use the spatial characteristics of the domain variables and predicts lithology solely based on geochemical characteristics. However, the continuity of the block lithology is implicitly enforced during the grade simulation stage by using soft boundaries for continuous grade variables. This results in the spatial connectivity being indirectly accounted for in the final lithology prediction.

A total of 100 realisations were used to assess the uncertainty in the estimated block domains. The most probable model for block lithology is determined by identifying the domain category with the highest probability for each block. The probability of a block belonging to a domain can be calculated as the proportion of predicted lithology out of the 100 realisations. Figure 12 illustrates the corresponding highest probability of the most likely domain for the cross-section shown in Figure 11. These maps quantify the uncertainty associated with the prediction of the most probable block lithology model.

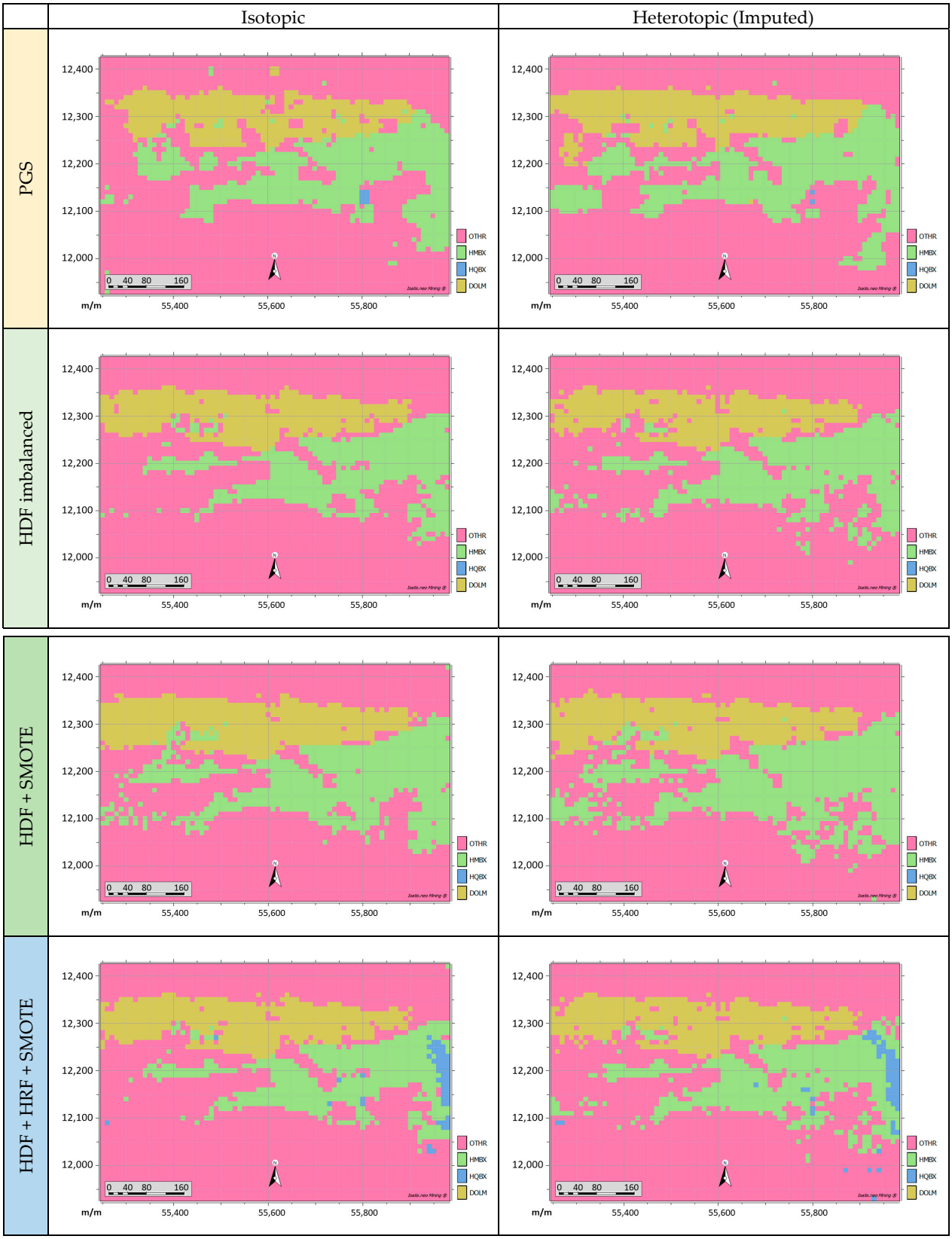

Figure 11. Top view of the horizontal cross-section for the most probable lithological model at an Elevation of Z = 9715 m.

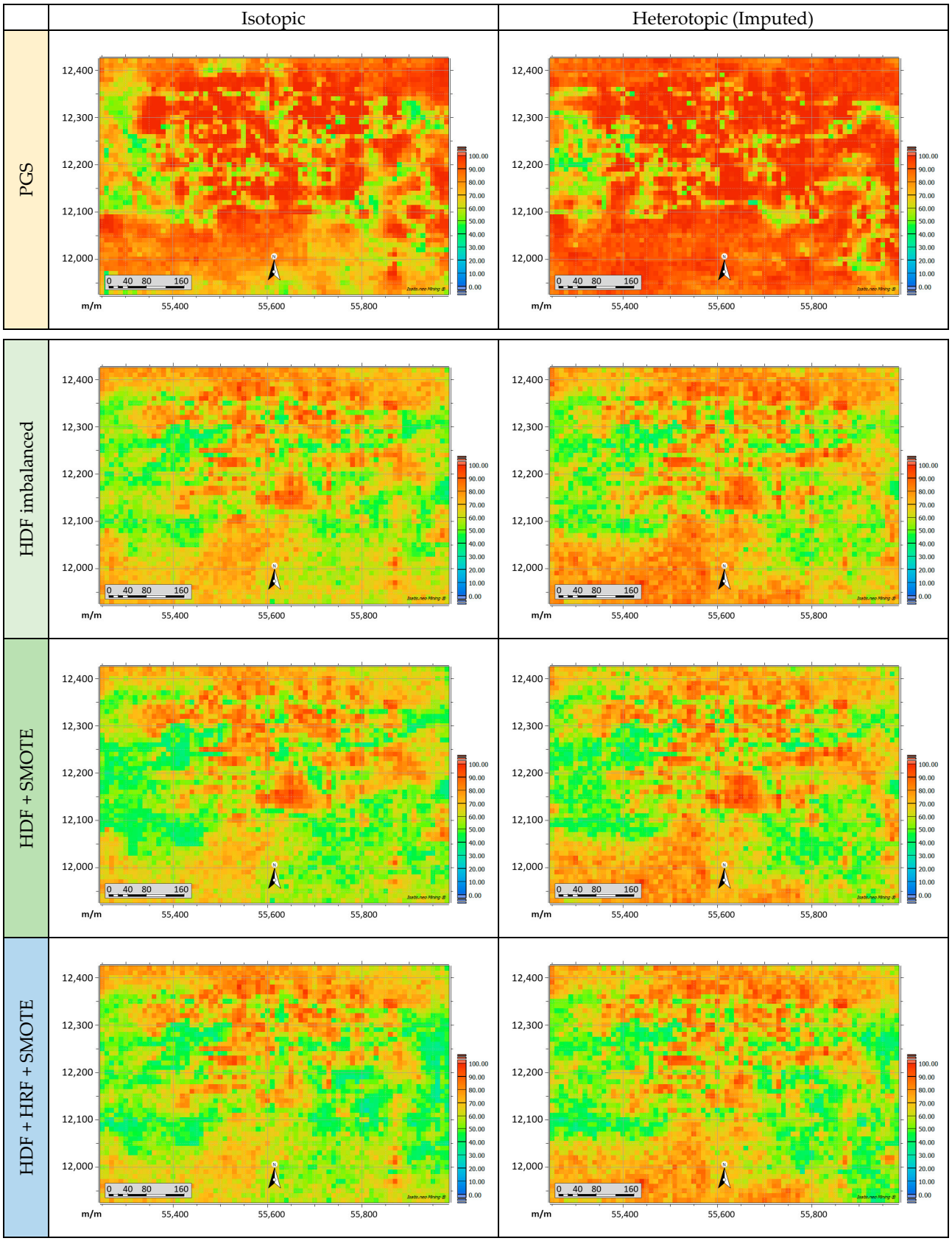

Figure 12. Top view of the horizontal cross-sections for the probability of the most probable domain at an elevation of Z = 9715 m.

3.4. Post-Processing the Realisations: Geometallurgical Domains

The use of realisation outcomes from HDF can assist in making informed decisions in a mining project. A significant benefit of this process is the ability to provide reliable domain boundaries with the corresponding probabilities, which will assist in correctly identifying mineral processing destinations for mined blocks. Simulation methods are used to estimate the uncertainty associated with domain boundaries and to identify areas with higher probability in relation to processing responses, such as NSR and grinding time. These areas can be regarded as geometallurgical domains that exhibit a certain level of confidence in decision-making processes related to mined blocks. Hence, this sub-section compares PGS and HDF in terms of two geometallurigcal models: NSR and grinding time.

3.4.1. Net Smelter Return

The NSR cut-off, which is used to determine the economic viability of a mining project, considers the revenue generated from copper and gold metals as well as the costs associated with mining operations, mineral processing, capital, and any penalty elements (in this case, fluorine and uranium). The NSR calculation also takes into account the estimated metallurgical recoveries of the minerals. In order to determine the NSR value for each block, metal recovery rates were estimated based on the expected sulphide mineralogy (for copper) or grade (for gold and silver) of each block. These recovery rates were derived using empirical models based on the processing plant performance. The block-by-block approach ensures that the NSR calculation accurately reflects the expected revenue generated from each individual block and helps to identify economically viable sections of the orebody. NSR values are calculated within the DOLM domain as it is most likely to contain gold and copper for the detailed analysis. The following simplified NSR formula was used to calculate the final results:

$$NSR_{simplified} = A \times Au_{ppm} + B \times Cu_{pct} - C \times F_{pct} - D \times U_{ppm}$$

where, A, B, C, and D are factors that depend on revenue, refining charges, metallurgical recovery assumptions, transport costs, and taxes. The most probable output for the DOLM domain was obtained from the PGS and HDF approaches chosen for the simulation of Au and Cu components and the deleterious F and U elements. Subsequently, the NSR is calculated using the following equation:

$$NSR_{simplified} = 1.15 \times Au_{ppm} + 29.4 \times Cu_{pct} - 0.51 \times F_{pct} - 0.0064 \times U_{ppm}$$

The E-type maps in Figures 13–15 represent the geometallurgical model for NSR values within the DOLM lithological domain over 100 conditional realisations, indicating areas with varying levels of certainty in finding particular geometallurgical domain boundaries. The blue sectors on the map represent areas with little profit, while the red sectors denote areas with a high level of potential profit. However, it is important to note that these descriptions are for average-case scenarios. The worst- and best-case scenarios can be found by taking into account the particular domain boundaries and grade realisations from the simulation results.

3.4.2. Grinding Time

One critical aspect of mining is the process of reducing the size of the extracted materials by grinding. Grinding is an energy-intensive process that accounts for a significant portion of the total energy consumption in mining operations. The amount of grinding required has a direct impact on energy efficiency, throughput, and the overall profitability of a mining operation. Thus, understanding the factors that contribute to grinding time is essential for optimising the process and improving the economic viability of mining operations.

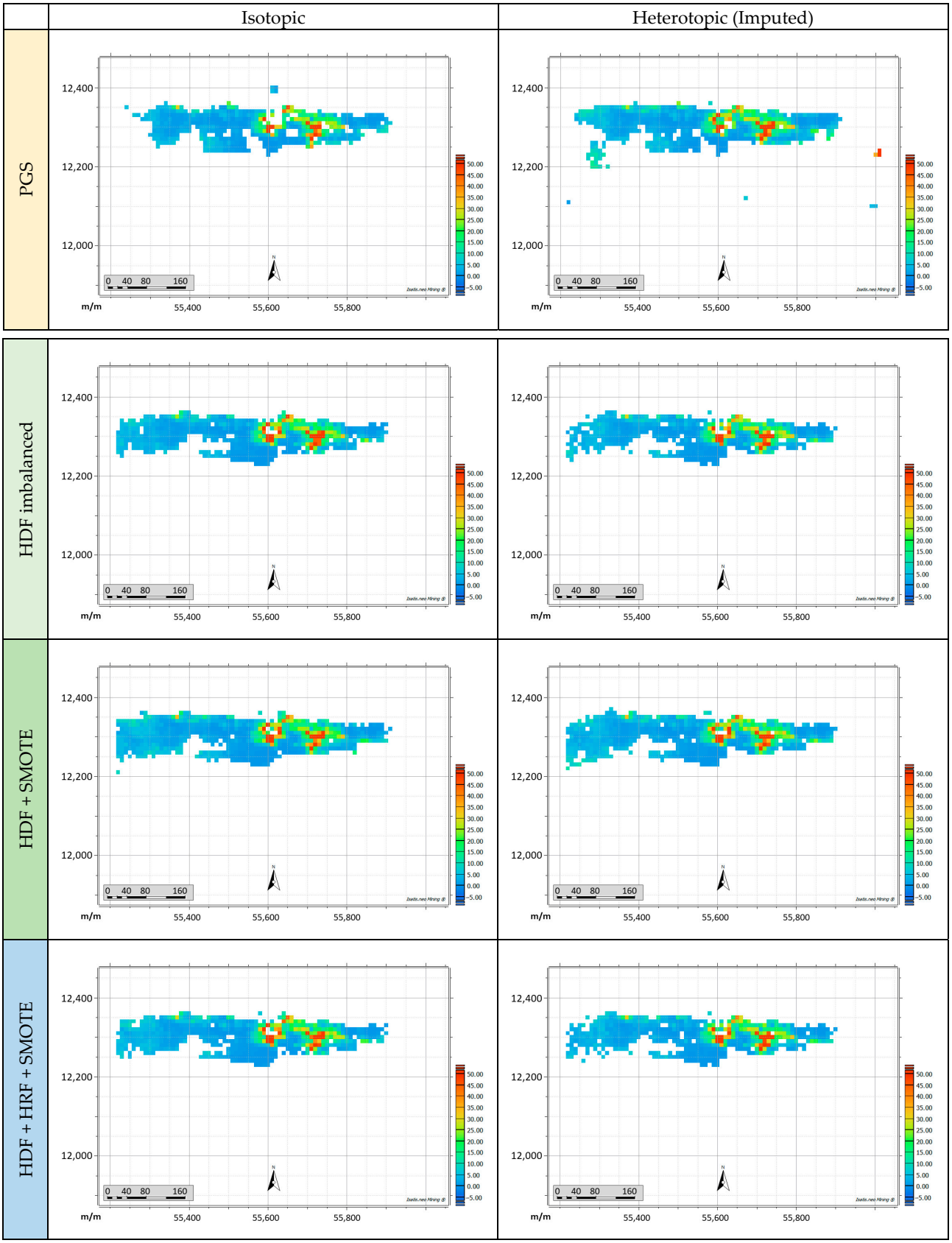

Figure 13. Top view of the horizontal cross-sections for the NSR with the most probable DOLM domain at an elevation of Z = 9715 m.

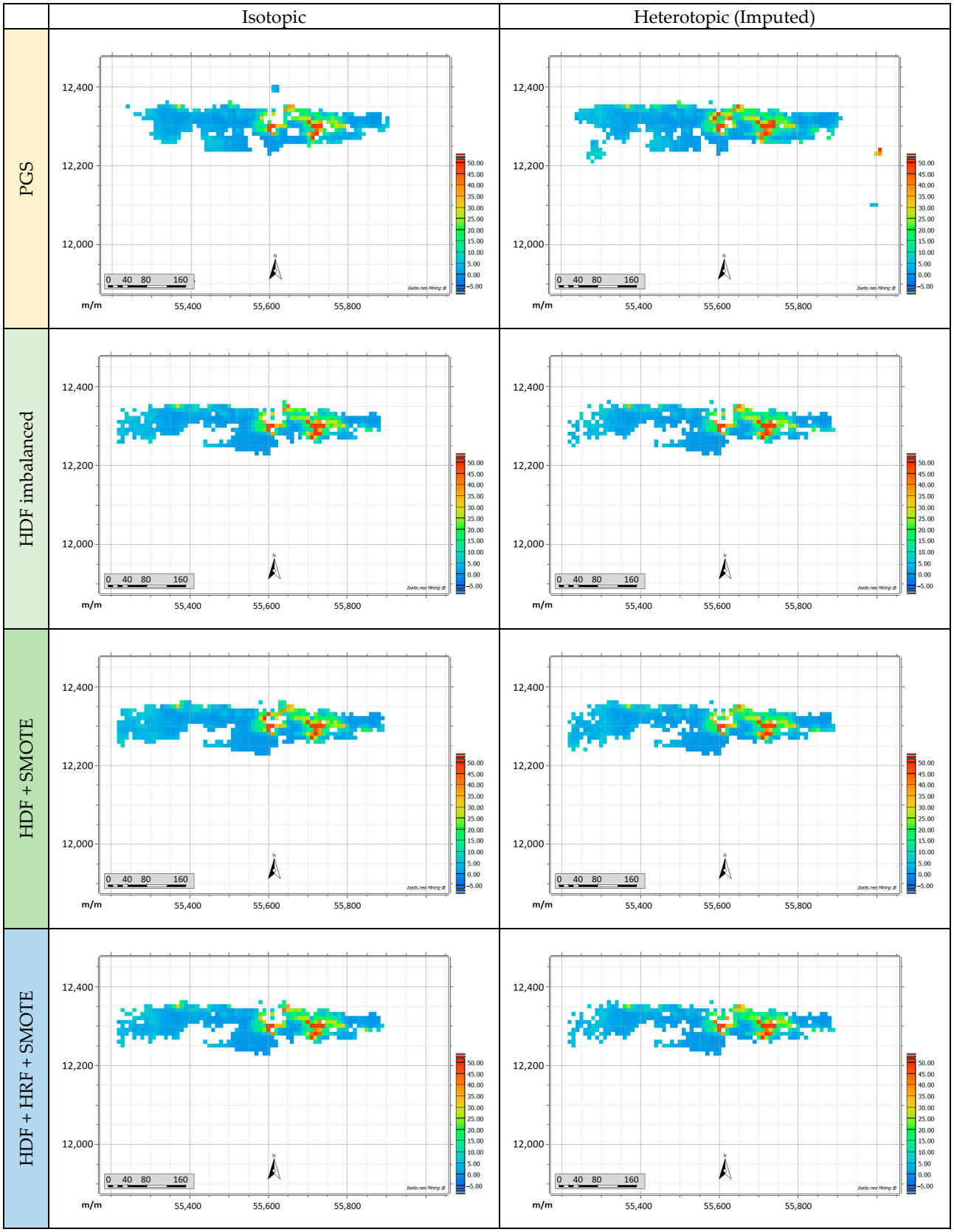

Figure 14. Top view of the horizontal cross-sections for the NSR with 50% probability of the DOLM domain at an elevation of Z = 9715 m.

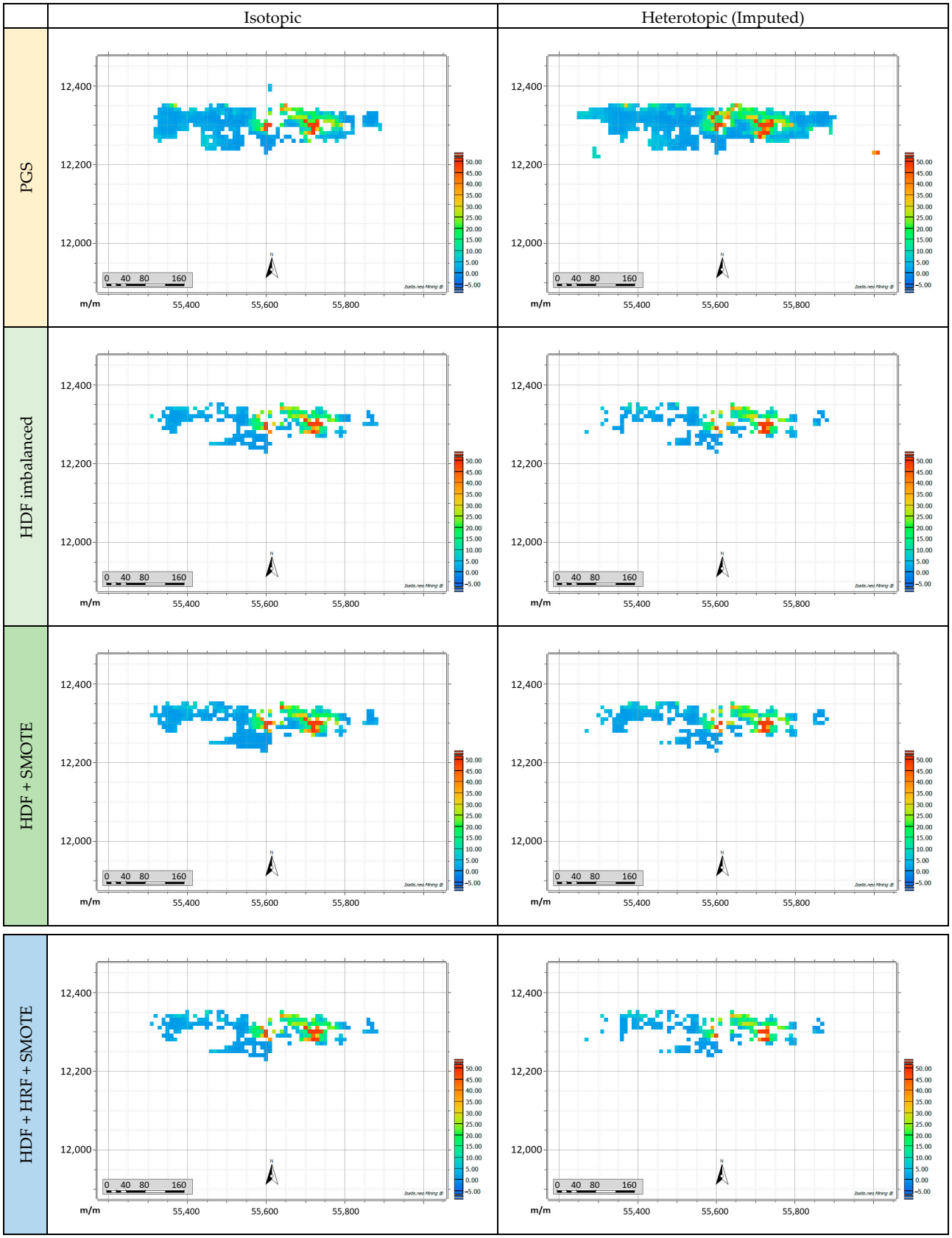

Figure 15. Top view of the horizontal cross-sections for the NSR with 70% probability of the DOLM domain at an elevation of Z = 9715 m.

The factors that influence grinding time include the ore type, the size distribution of the crushed ore, the hardness of the ore, the grinding process, and the operating conditions. These factors can impact the grinding efficiency, which is typically measured by the amount of material that can be ground per unit of energy consumed. In this case study, HMBX and HQBX domains have significant differences in grinding time, although it is easy to misclassify these two lithologies during the core logging process. The HQBX composited ore requires 1.5–1.8 times more time to grind than ore with HMBX content. Thus, the prediction of the boundaries of these two domains is critical, as the lithological model, especially the HQBX domain, is directly related to geometallurigcal responses such as grinding time.

In the most probable maps shown above, differences in the presence of HQBX and HMBX may vary depending on each domaining method. As a result, the grinding time for particular blocks may differ. Assuming that the general conditions are similar for HQBX and HMBX, a comparison in terms of the number of blocks per domain shows the extent to which the grinding time can change from one method to another (Table 10). As a result, noise filtering significantly affects the final outcome for the HQBX class, and thus grinding time is likely to be longer for some of the mining blocks. Consequently, decisions related to processing require revision.

Table 10. Number of blocks for HQBX and HMBX classified by probability and domaining methods (total number of blocks is 260,000).

		HMBX				HQBX			
	Methods	**Most Probable**	**Prob. >50**	**Prob. >60**	**Prob. >70**	**Most Probable**	**Prob. >50**	**Prob. >60**	**Prob. >70**
Isotopic	PGS	58,569	48,090	37,575	29,420	1337	966	673	441
	HDF imbalanced	48,798	32,743	17,511	7871	0	0	0	0
	HDF + SMOTE	68,881	43,586	23,961	11,667	3	0	0	0
	HDF + HRF + SMOTE	39,974	17,982	8176	2753	4945	737	168	20
Heterotopic	PGS	44,272	41,819	35,324	28,943	2261	2111	1421	929
	HDF imbalanced	51,187	40,484	23,317	11,362	0	0	0	0
	HDF + SMOTE	66,173	50,211	30,477	15,702	2	0	0	0
	HDF + HRF + SMOTE	43,263	24,320	12,306	5118	6286	1615	550	145

4. Results and Discussion

The proposed domaining framework includes the HybridRepairFilter noise filtering method, which has resulted in a significant change in the domain layout. The filtering involves the removal of 2029 samples and the re-labelling of 1953 samples from the original HMBX class to the HQBX class. When the preprocessing is implemented, there is a significant increase in the number of blocks classified as HQBX, as shown in Table 11. Without preprocessing, none of the blocks were classified as HQBX lithology. Moreover, the singular balancing by SMOTE, excluding noise filtering, resulted in 2 or 3 HQBX blocks for the isotopic or heterotopic data. The number of HQBX blocks increased to 4945 for the isotopic subset and to 6286 for the heterotopic subset when noise filtering preprocessing was used. The use of the multiclass HRF noise filtering method resulted in significant changes in the HMBX and HQBX domains. Other domains have slight changes, which can be seen in the most probable results in Table 11.

The proposed domaining framework may be impacted by three main sources of uncertainty related to the domain boundary definition. The first source is the inherent uncertainty associated with the data, which can result in significant errors due to mislogging. While data uncertainty cannot be completely eliminated, it can be reduced through noise filtering for important domains such as HMBX and HQBX. The second source of uncertainty comes from the simulation results, which can be used to generate multiple realisations for uncertainty quantification. Finally, there is model's uncertainty associated with the classification algorithm and the SMOTE balancing method, which relies heavily on input data to produce reliable outputs. This classification model uncertainty

is directly influenced by the previous two sources of uncertainty. As a consequence, the proposed domaining framework can be impacted by multiple sources of uncertainty, and reducing these impacts requires each of these sources to be assessed. In this case, the probability maps can be used to mitigate the risk associated with the overall uncertainty in the generated domain boundaries. Using thresholds, blocks with high certainty can be selected for further resource classification.

Table 11. Statistics for the most probable results for OTHR, HQBX, HMBX and DOLM domains for the various methods.

		Isotopic				Heterotopic (Imputed)			
		OTHR	**HQBX**	**HMBX**	**DOLM**	**OTHR**	**HQBX**	**HMBX**	**DOLM**
PGS	Number of blocks	183,092	1337	58,569	17,002	194,283	2261	44,272	19,184
	Proportion (%)	70.42	0.51	22.53	6.54	74.72	0.87	17.03	7.38
HDF imbalanced	Number of blocks	196,240	0	48,798	14,962	195,644	0	51,187	13,169
	Proportion (%)	75.48	0.00	18.77	5.75	75.25	0.00	19.69	5.07
HDF + SMOTE	Number of blocks	172,416	3	68,881	18,700	177,596	2	66,173	16,229
	Proportion (%)	66.31	0.00	26.49	7.19	68.31	0.00	25.45	6.24
HDF + HRF + SMOTE	Number of blocks	199,768	4945	39,974	15,313	197,144	6286	43,263	13,307
	Proportion (%)	76.83	1.91	15.37	5.89	75.82	2.42	16.64	5.12

4.1. Geometallurgical Model: Net Smelter Return

A NSR cut-off and specific mining dimensions may be used to optimise stope layouts. The optimisation process is informed by the local orientation of the interpreted mineralization zone, with high grades commonly represented using wireframes (i.e., triangulations). Consequently, some material below the specified cut-off grade may be included, and some material above the cut-off value may be excluded during future extraction. In this study, the calculated NSR is relative to the original NSR used on the mine site, which is based on the revenue generated from the extracted metals and the expenses incurred in site operations. The NSR was computed for the DOLM domain produced by PGS and HDF with different configurations.

In the maps shown in Section 3.4.1, the general DOLM domain boundaries are similar for all the methods. However, the probabilities of PGS and HDF are significantly different because the DOLM boundaries shrink rapidly as the threshold for the domain probability approaches its maximum. For example, when the DOLM domain for predicting NSR is set to a probability of 70% (P-70), the number of HDF blocks decreases to more than twice that with a probability of 50% (P-50), whereas the number of blocks for PGS decreases by about 1.5 times (Table 12). The overall mean values for NSR stay relatively stable for all the corresponding thresholds (cut-offs), which is expected for the DOLM class because of its distinct geochemical characteristics. Consequently, HDF and PGS produce similar results for the geometallurgical model of NSR values.

Table 12. Global characteristics of NSR within the DOLM domain for different domaining methods and corresponding probabilities (MP—most probable, P-50—probability over 50%, P-70—probability over 70%).

	Methods	**Threshold**	**Number of Blocks**	**Mean**	**Std Dev.**	**Q5**	**Q25**	**Q50**	**Q75**	**Q95**
Isotopic	PGS	MP	17,002	7.17	11.40	0.40	1.48	3.92	8.53	23.47
		P-50	16,813	7.03	11.22	0.40	1.44	3.88	8.34	22.79
		P-70	10,691	6.36	11.78	0.36	1.11	2.69	6.53	24.05
	HDF imbalanced	MP	14,962	7.00	12.80	0.35	1.16	2.92	7.13	26.91
		P-50	12,295	6.54	12.52	0.32	1.01	2.49	6.09	26.65
		P-70	4610	6.66	12.61	0.30	0.83	1.97	6.25	30.32

Table 12. *Cont.*

	Methods	Threshold	Number of Blocks	Mean	Std Dev.	Q5	Q25	Q50	Q75	Q95
	HDF + SMOTE	MP	18,700	7.02	12.26	0.38	1.32	3.42	7.66	24.14
		P-50	14,531	6.55	12.15	0.34	1.07	2.68	6.58	25.12
		P-70	5966	6.42	12.38	0.29	0.85	2.00	5.68	29.75
	HDF + HRF + SMOTE	MP	15,313	7.21	13.06	0.34	1.18	3.03	7.35	27.42
		P-50	12,374	6.51	12.45	0.32	1.01	2.48	6.15	26.31
		P-70	4594	6.56	12.34	0.29	0.83	1.98	5.87	30.13
Heterotopic	PGS	MP	19,184	7.76	12.79	0.40	1.45	3.81	8.78	27.54
		P-50	18,045	7.63	12.91	0.39	1.40	3.62	8.32	27.69
		P-70	14,039	7.21	13.16	0.37	1.22	3.04	7.22	28.05
	HDF imbalanced	MP	13,169	6.94	13.23	0.34	1.08	2.73	6.56	27.77
		P-50	11,042	6.45	12.42	0.31	0.95	2.36	5.94	27.19
		P-70	4027	6.64	12.15	0.28	0.81	2.06	6.64	29.91
	HDF + SMOTE	MP	16,229	6.94	12.70	0.36	1.19	3.07	7.12	25.96
		P-50	13,089	6.43	12.27	0.33	1.02	2.55	6.04	26.46
		P-70	5315	6.49	12.25	0.29	0.84	2.03	5.93	28.75
	HDF + HRF + SMOTE	MP	13,307	6.97	13.19	0.34	1.10	2.73	6.59	27.75
		P-50	11,160	6.44	12.50	0.31	0.97	2.38	5.98	27.59
		P-70	4022	6.76	12.42	0.29	0.82	2.09	6.46	28.85

4.2. Geometallurgical Model: Grinding Time—HQBX

The grinding times for HMBX and HQBX may differ due to the differences in their physical and chemical properties. HQBX contains quartz in addition to hematite, while HMBX contains only hematite. Quartz is a harder mineral than hematite, which means that Hematite Quartz Breccia requires a longer grinding time to achieve the desired particle size compared with Hematite Breccia.

According to the mine geologists and actual processing (i.e., grinding) results, it takes around 5–6 min for a unit of HMBX and 9–11 min for a unit of HQBX to be ground to the desired particle size and quality. The number of blocks for HQBX varies significantly for the PGS and HDF domaining methods. In Table 13, the results for PGS reproduce the original HQBX proportions for isotopic and heterotopic data. As a result, there is a good HQBX presence. However, classifiers in the case of HDF with imbalanced and balanced (i.e., SMOTE) completely disregard the HQBX in their prediction. The reasons behind this are the low number of HQBX samples, the inherent noise in the data, and the similarity of the HQBX and HMBX classes. On the other hand, HRF solves the problem of noisy data and provides distinct statistical boundaries for the classifier [30]. Consequently, the balanced HDF with HRF predicts many more blocks than PGS because it depends on the simulated grade values (Figure 16) and not directly on the domain proportions of samples. This is advantageous for irregularly spaced data, such as in this case study.

Table 13. Number of blocks for the HQBX domain classified by probability and domaining methods (total number of blocks: 260,000).

Methods	Isotopic				Heterotopic			
	Most Probable	Prob. >50	Prob. >60	Prob. >70	Most Probable	Prob. >50	Prob. >60	Prob. >70
PGS	1337	966	673	441	2261	2111	1421	929
HDF imbalanced	0	0	0	0	0	0	0	0
HDF + SMOTE	3	0	0	0	2	0	0	0
HDF + HRF + SMOTE	4945	737	168	20	6286	1615	550	145

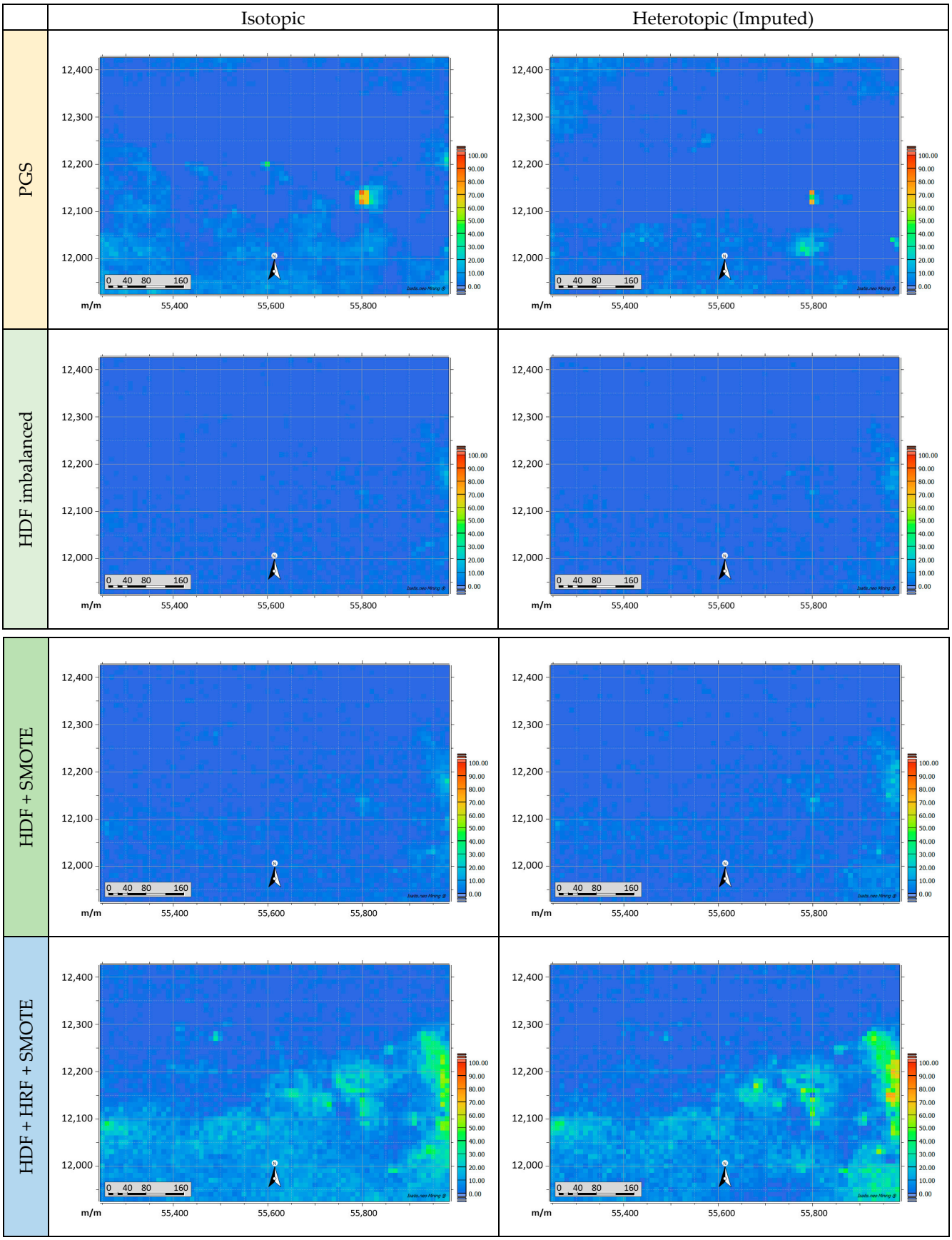

Figure 16. Top view of the horizontal cross-sections for HQBX probability at an elevation of Z = 9715 m.

4.3. General Performance

The outcomes derived from the proposed HDF domaining framework demonstrate considerable potential, particularly in situations that involve unbalanced datasets. The framework can accommodate a higher number of domains and can differentiate domains that appear to be almost identical classes/domains, such as HMBX and HQBX, by using noise-filtering approaches. Abildin et al. [30] demonstrated the application of the HDF for a multiclass problem, but with the HRF applied for the binary class. In this paper, HRF is used for the multiclass problem using the pairwise approach, in which noise for each class is filtered out against all other classes. A comparison of the most probable PGS and HDF maps for the DOLM domain shows that the domain boundaries are similar, although the probabilities are different. In the case of the HQBX domain, the reproduction of the class is different, but the most probable maps show similar features in the same locations. Thus, the geometallurgical model for NSR is similar to that of PGS and the modified HDF. However, the reproduction of the geometallurgical model for HQBX is better for the HDF, as demonstrated in the validation in Abildin et al. [30]. Overall, for stable domains, such as DOLM, the geometallurgical models produced by the modified HDF are similar to the models obtained by benchmarking PGS. HDF can handle complex geometallurgical domains such as HQBX by including quantitated covariates.

The use of HDF provides the added advantage of quantifying the level of uncertainty associated with the domain model, as can be seen in the analyses provided in this paper. Moreover, it incorporates more abundant grade values into the domaining process, and thus, the reliability of the domain boundaries increases. HRF and SMOTE mitigate data problems such as noise and imbalance. The benefit of using all important information significantly impacts the evaluation of confidence in the resource model as it enables more accurate classification of resource categories based on the JORC code [55].

5. Conclusions

The accurate definition of domains is critical for the estimation of mineral resources. The current practises for domaining include labour-intensive explicit modelling, complex implicit modelling, and tedious geostatistical methods such as PluriGaussian simulation. To address these challenges, the authors previously proposed a hybrid data-driven machine learning-based domaining framework (HDF) that combines geostatistical simulations and classification algorithms with noise filtering preprocessing (HRF). HDF is data-driven, using only grade data and lithology logs to define domain boundaries. In this work, the method has been compared with PGS, which simulates categorical variables by using the truncation rule on Gaussian random fields.

HDF and PGS have been tested using a dataset from an IOCG deposit that has complex geological features. The proposed domaining method uses the turning bands method for simulations, an extreme gradient boosting classification algorithm for classification, and an ensemble of machine learning methods for noise removal for the multiclass problem and detection of potentially miss-logged domain categories. The preprocessing of data has been demonstrated to enhance the performance of the classification algorithm, leading to better results produced by the domaining framework. The proposed method generates a block-domain model with multiple realisations, from which the most likely block-domain model and the corresponding probability can be obtained to quantify uncertainty. The impacts of the produced domain boundaries on the geometallurgical models, such as NSR and grinding time, are further examined in this work.

HDF is a versatile approach that can be applied to other datasets originating from diverse deposits and is not restricted to the modelling of lithological variables. This framework can be extended to model any categorical variable that depends on continuous variables, such as the geometallurgical domains investigated in this study. By using this method, it is anticipated that the accuracy of mineral resource estimation will be enhanced, resulting in increased confidence in the location and quantity of valuable minerals.

Author Contributions: Y.A.: Conceptualisation, Methodology, Software, Validation, Investigation, Writing—Original Draft, Writing—Review & Editing, Visualisation. C.X., P.D. and A.A.: Conceptualisation, Methodology, Resources, Writing—Review & Editing, Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been supported by the South Australian Government through the PRIF RCP Mining Consortium "Unlocking Complex Resources through Lean Processing".

Data Availability Statement: Data is not allowed to share since it has sensitive information.

Acknowledgments: The authors thank the Geovariances Company for providing their software for academic use.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Coward, S.; Dowd, P. Geometallurgical models for the quantification of uncertainty in mining project value chains. In Proceedings of the International Symposium on the Application of Computers and Operations Research in the Mineral Industry (APCOM), Fairbanks, AK, USA, 23–27 May 2015; pp. 360–369.
2. Macfarlane, A.; Williams, T. Optimizing value on a copper mine by adopting a geometallurgical solution. *J. South. Afr. Inst. Min. Metall.* **2014**, *114*, 929–935.
3. Tercan, A.; Sohrabian, B. Multivariate geostatistical simulation of coal quality data by independent components. *Int. J. Coal Geol.* **2013**, *112*, 53–66. [CrossRef]
4. Dowd, P.; Xu, C.; Coward, S. Strategic mine planning and design: Some challenges and strategies for addressing them. *Min. Technol.* **2016**, *125*, 22–34. [CrossRef]
5. Lishchuk, V.; Koch, P.-H.; Ghorbani, Y.; Butcher, A.R. Towards integrated geometallurgical approach: Critical review of current practices and future trends. *Miner. Eng.* **2020**, *145*, 106072. [CrossRef]
6. Deutsch, J.; Szymanski, J.; Etsell, T. Metallurgical variable re-expression for geostatistics. In *Geostatistical and Geospatial Approaches for the Characterization of Natural Resources in the Environment: Challenges, Processes and Strategies*; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 83–88.
7. Abildin, Y.; Madani, N.; Topal, E. A hybrid approach for joint simulation of geometallurgical variables with inequality constraint. *Minerals* **2019**, *9*, 24. [CrossRef]
8. 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]
9. Sepulveda, E.; Dowd, P.; Xu, C.; Addo, E. Multivariate modelling of geometallurgical variables by projection pursuit. *Math. Geosci.* **2017**, *49*, 121–143. [CrossRef]
10. Cowan, E.; Beatson, R.; Ross, H.; Fright, W.; McLennan, T.; Evans, T.; Carr, J.; Lane, R.; Bright, D.; Gillman, A. Practical implicit geological modelling. In Proceedings of the 5th International Mining Geology Conference, Bendigo, VIC, Australia, 17–19 November 2003; pp. 89–99.
11. Gonçalves, Í.G.; Kumaira, S.; Guadagnin, F. A machine learning approach to the potential-field method for implicit modeling of geological structures. *Comput. Geosci.* **2017**, *103*, 173–182. [CrossRef]
12. Lajaunie, C.; Courrioux, G.; Manuel, L. Foliation fields and 3D cartography in geology: Principles of a method based on potential interpolation. *Math. Geol.* **1997**, *29*, 571–584. [CrossRef]
13. Manchuk, J.G.; Deutsch, C.V. Boundary modeling with moving least squares. *Comput. Geosci.* **2019**, *126*, 96–106. [CrossRef]
14. Armstrong, M.; Galli, A.; Beucher, H.; Loc'h, G.; Renard, D.; Doligez, B.; Eschard, R.; Geffroy, F. *Plurigaussian Simulations in Geosciences*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011.
15. Emery, X.; Séguret, S.A. *Geostatistics for the Mining Industry: Applications to Porphyry Copper Deposits*; CRC Press: Boca Raton, FL, USA, 2020.
16. Journel, A.; Isaaks, E. Conditional indicator simulation: Application to a Saskatchewan uranium deposit. *J. Int. Assoc. Math. Geol.* **1984**, *16*, 685–718. [CrossRef]
17. Journel, A.G. Nonparametric estimation of spatial distributions. *J. Int. Assoc. Math. Geol.* **1983**, *15*, 445–468. [CrossRef]
18. Xu, C.; Dowd, P.A.; Mardia, K.V.; Fowell, R.J. A flexible true plurigaussian code for spatial facies simulations. *Comput. Geosci.* **2006**, *32*, 1629–1645. [CrossRef]
19. Emery, X.; Ortiz, J.; Cáceres, A. Geostatistical modelling of rock type domains with spatially varying proportions: Application to a porphyry copper deposit. *J. South. Afr. Inst. Min. Metall.* **2008**, *108*, 284–292.
20. Maleki, M.; Emery, X.; Cáceres, A.; Ribeiro, D.; Cunha, E. Quantifying the uncertainty in the spatial layout of rock type domains in an iron ore deposit. *Comput. Geosci.* **2016**, *20*, 1013–1028. [CrossRef]
21. Silva, D. Enhanced Geologic Modeling of Multiple Categorical Variables. Ph.D. Thesis, University of Alberta, Edmonton, AB, Canada, 2018.

22. Séguret, S.A. Analysis and estimation of multi-unit deposits: Application to a porphyry copper deposit. *Math. Geosci.* **2013**, *45*, 927–947. [CrossRef]
23. Adeli, A.; Emery, X.; Dowd, P. Geological modelling and validation of geological interpretations via simulation and classification of quantitative covariates. *Minerals* **2017**, *8*, 7. [CrossRef]
24. Fouedjio, F.; Hill, E.J.; Laukamp, C. Geostatistical clustering as an aid for ore body domaining: Case study at the Rocklea Dome channel iron ore deposit, Western Australia. *Appl. Earth Sci.* **2018**, *127*, 15–29. [CrossRef]
25. Madani, N.; Maleki, M.; Emery, X. Nonparametric geostatistical simulation of subsurface facies: Tools for validating the reproduction of, and uncertainty in, facies geometry. *Nat. Resour. Res.* **2019**, *28*, 1163–1182. [CrossRef]
26. Moreira, G.d.C.; Coimbra Leite Costa, J.F.; Marques, D.M. Defining geologic domains using cluster analysis and indicator correlograms: A phosphate-titanium case study. *Appl. Earth Sci.* **2020**, *129*, 176–190. [CrossRef]
27. Sepúlveda, E.; Dowd, P.; Xu, C. Fuzzy clustering with spatial correction and its application to geometallurgical domaining. *Math. Geosci.* **2018**, *50*, 895–928. [CrossRef]
28. Kasmaee, S.; Raspa, G.; de Fouquet, C.; Tinti, F.; Bonduà, S.; Bruno, R. Geostatistical estimation of multi-domain deposits with transitional boundaries: A sensitivity study for the Sechahun iron mine. *Minerals* **2019**, *9*, 115. [CrossRef]
29. Amarante, F.A.N.; Rolo, R.M.; Coimbra Leite Costa, J.F. Boundary simulation–a hierarchical approach for multiple categories. *Appl. Earth Sci.* **2021**, *130*, 114–130. [CrossRef]
30. Abildin, Y.; Xu, C.; Dowd, P.; Adeli, A. A hybrid framework for modelling domains using quantitative covariates. *Appl. Comput. Geosci.* **2022**, *16*, 100107. [CrossRef]
31. Adeli, A.; Emery, X. Geostatistical simulation of rock physical and geochemical properties with spatial filtering and its application to predictive geological mapping. *J. Geochem. Explor.* **2021**, *220*, 106661. [CrossRef]
32. Rossi, M.E.; Deutsch, C.V. *Mineral Resource Estimation*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
33. Matheron, G.; Beucher, H.; de Fouquet, C.; Galli, A.; Guérillot, D.; Ravenne, C. Conditional simulation of the geometry of fluvio-deltaic reservoirs. In Proceedings of the SPE Annual Technical Conference and Exhibition, Dallas, TX, USA, 27–30 September 1987.
34. Sadeghi, M.; Madani, N.; Falahat, R.; Sabeti, H.; Amini, N. Hierarchical reservoir lithofacies and acoustic impedance simulation: Application to an oil field in SW of Iran. *J. Pet. Sci. Eng.* **2022**, *208*, 109552. [CrossRef]
35. Dowd, P.; Pardo-Igúzquiza, E.; Xu, C. Plurigau: A computer program for simulating spatial facies using the truncated plurigaussian method. *Comput. Geosci.* **2003**, *29*, 123–141. [CrossRef]
36. Madani, N.; Emery, X. Plurigaussian modeling of geological domains based on the truncation of non-stationary Gaussian random fields. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 893–913. [CrossRef]
37. Emery, X.; Lantuéjoul, C. Tbsim: A computer program for conditional simulation of three-dimensional gaussian random fields via the turning bands method. *Comput. Geosci.* **2006**, *32*, 1615–1628. [CrossRef]
38. Le Loc'h, G.; Galli, A. Truncated plurigaussian method: Theoretical and practical points of view. *Geostat. Wollongong* **1997**, *96*, 211–222.
39. Emery, X. Simulation of geological domains using the plurigaussian model: New developments and computer programs. *Comput. Geosci.* **2007**, *33*, 1189–1201. [CrossRef]
40. Lantuéjoul, C. *Geostatistical Simulation: Models and Algorithms*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2001.
41. Arroyo, D.; Emery, X. Spectral simulation of vector random fields with stationary Gaussian increments in d-dimensional Euclidean spaces. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 1583–1592. [CrossRef]
42. Emery, X. A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields. *Comput. Geosci.* **2008**, *34*, 1850–1862. [CrossRef]
43. Friedman, J.H.; Tukey, J.W. A projection pursuit algorithm for exploratory data analysis. *IEEE Trans. Comput.* **1974**, *100*, 881–890. [CrossRef]
44. Barnett, R.M.; Manchuk, J.G.; Deutsch, C.V. Projection pursuit multivariate transform. *Math. Geosci.* **2014**, *46*, 337–359. [CrossRef]
45. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
46. Gu, Y.; Zhang, D.; Bao, Z. Lithological classification via an improved extreme gradient boosting: A demonstration of the Chang 4+ 5 member, Ordos Basin, Northern China. *J. Asian Earth Sci.* **2021**, *215*, 104798. [CrossRef]
47. Morales, P.; Luengo, J.; Garcia, L.P.F.; Lorena, A.C.; de Carvalho, A.C.; Herrera, F. The NoiseFiltersR Package: Label Noise Preprocessing in R. *R J.* **2017**, *9*, 219. [CrossRef]
48. Miranda, A.L.; Garcia, L.P.F.; Carvalho, A.C.; Lorena, A.C. Use of classification algorithms in noise detection and elimination. In Proceedings of the Hybrid Artificial Intelligence Systems: 4th International Conference, HAIS 2009, Salamanca, Spain, 10–12 June 2009; pp. 417–424.
49. Fernández, A.; Garcia, S.; Herrera, F.; Chawla, N.V. SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *J. Artif. Intell. Res.* **2018**, *61*, 863–905. [CrossRef]
50. Williams, P.J.; Barton, M.D.; Johnson, D.A.; Fontboté, L.; De Haller, A.; Mark, G.; Oliver, N.H.; Marschik, R. Iron oxide copper-gold deposits: Geology, space-time distribution, and possible modes of origin. *Econ. Geol.* **2005**, *100*, 371–405. [CrossRef]
51. Belperio, A.; Flint, R.; Freeman, H. Prominent Hill: A hematite-dominated, iron oxide copper-gold system. *Econ. Geol.* **2007**, *102*, 1499–1510. [CrossRef]

52. Stekhoven, D.J.; Bühlmann, P. MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics* **2012**, *28*, 112–118. [CrossRef] [PubMed]
53. Emery, X. Testing the correctness of the sequential algorithm for simulating Gaussian random fields. *Stoch. Environ. Res. Risk Assess.* **2004**, *18*, 401–413. [CrossRef]
54. Leuangthong, O.; McLennan, J.A.; Deutsch, C.V. Minimum acceptance criteria for geostatistical realizations. *Nat. Resour. Res.* **2004**, *13*, 131–141. [CrossRef]
55. Code, J. Australasian code for reporting of exploration results, mineral resources and ore reserves. *AusIMM Melb.* **2012**, *44*, 320.