The Robert M. Buchan Department of Mining

Predictive Geometallurgy and Geostatistics Lab

Annual Report 2022

Edited by: Julian M. Ortiz

Queen's University

Predictive Geometallurgy and Geostatistics Lab
Queen's University

Annual report 2022

This report summarizes the ongoing research of the Predictive Geometallurgy and Geostatistics Laboratory at Queen's University in Kingston, Ontario, Canada. 2021 was a challenging year, with continuous restrictions to meetings and limited group interactions. Despite these difficulties, the lab completed important and novel work. The delay in the release of this report is just another sign of the strain of the last couple of years, however, results are encouraging and the lab is doing important contributions to research and industry.

This year, two students graduated, one Master of Applied Science and one Doctor of Philosophy, two new students joined the group. The following two theses were completed in this period:

- Kasimcan Koruk, M.A.Sc. (Aug. 2022), "***Definition of geological domains with an ensemble implementation of Support Vector Classification***"
- Alvaro Riquelme, Ph.D. (Sep. 2022), "***Multivariate simulation using a locally varying coregionalization model***"

The hold on Sebastian Avalos Ph.D. thesis has been lifted after the successful filing of a patent:

- Sebastian Avalos, Ph.D. (Sep. 2021), "***Advanced predictive methods applied to geometallurgical modelling***"

The research group is currently composed of:

- Sebastian Avalos, Post-Doc
- David Casson, Ph.D. student
- Soheil Kheirparast, M.A.Sc. student
- Paula Larrondo, Ph.D. student
- Tong Li, Visiting Ph.D. student
- Alvaro Mariño, M.A.Sc. student
- Noble Potakey, M.A.Sc. student
- Carlos Moraga, Ph.D. student

We continued collaboration with other faculty members and researchers, including:

- Asli Sari, Assistant Professor – The Robert M. Buchan Department of Mining (Queen's University). Dr. Sari and Dr. Ortiz co-supervised Soheil Kheirparast in his M.A.Sc.
- Raimon Tolosana-Delgado, Senior Scientist (Helmholtz-Zentrum Dresden-Rossendorf). Dr. Tolosana-Delgado hosted a research internship of Sebastian Avalos.

- Brian Frank, Professor – Electrical and Computer Engineering (Queen's University). Dr. Frank and Dr. Ortiz co-supervise Paula Larrondo in her Ph.D.

Nine contributions are available this year, totaling 140 pages, with reviews on topics such as clustering methods, grade control, models for muck pile blast movement, generative adversarial networks (GANs), and interesting progress in non-Gaussian models, mineral prospectivity and bulk ore sorting. Industrial collaboration continues with SRK Consulting Canada and MineSense.

As always, we welcome industrial and academic collaboration. This provides opportunities to fund new graduate students and novel research, and directly benefits industrial partners. If interested, please send a note to julian.ortiz@queensu.ca.

Julian M. Ortiz
Associate Professor, The Robert M. Buchan Department of Mining
Director, Predictive Geometallurgy and Geostatistics Lab
Queen's University

December 2022

Table of contents

Journal and Conference Publications and Presentations

Publications in book chapters, peer-reviewed journals and international conferences are listed below for 2021. These are not included in this report, since the copyright belongs to the corresponding publishers, but can be requested for personal use or research purposes directly to julian.ortiz@queensu.ca.

Patent

1. Avalos SA, Ortiz JM (2022) ***Mine Scheduling Methods and Constructs***, Provisional Patent US 63/388,016, filed 07/19/2022.

Book chapters

1. Ortiz JM, Silva JF (2022) ***Entropy***, in Encyclopedia of Mathematical Geosciences, Daya Sagar B, Cheng Q, McKinley J, Agterberg F (Eds.), Encyclopedia of Earth Sciences Series, Springer, Cham, 5 p. https://doi.org/10.1007/978-3-030-26050-7_102-1

Journal papers

1. Moraga C, Kracht W, Ortiz JM (2022) ***Process simulation to determine blending and residence time distribution in mineral processing plants***, Minerals Engineering, 187: 107807. https://doi.org/10.1016/j.mineng.2022.107807
2. Jelvez E, Morales N, Ortiz JM (2022) ***Stochastic final pit limits: an efficient frontier analysis under geological uncertainty in the open-pit mining industry***, Mathematics, 10(1): 100. https://doi.org/10.3390/math10010100

Conference papers and presentations

1. Ortiz JM (2022) ***Predictive Modelling in Geometallurgy***, in Geomet-Procemin 2022, Santiago, October 5-7 2022.
2. Avalos S, Ortiz JM (2022) ***Spatial multivariate morphing transformation applied to geometallurgical attributes***, in Geomet-Procemin 2022, Santiago, October 5-7 2022.
3. Riquelme A, Ortiz JM (2022) ***Simulation of complex multivariate relationships based on a non-stationary coregionalization model***, in 21st Annual Conference of the International Association for Mathematical Geosciences – IAMG 2022, Nancy, France, Aug 29-Sep 3, 2022.
4. Koruk K, Ortiz JM (2022) ***Definition of geological domains with Ensemble Support Vector Classification***, in 21st Annual Conference of the International Association for Mathematical Geosciences – IAMG 2022, Nancy, France, Aug 29-Sep 3, 2022.

5. Ortiz JM, Avalos S, Frenzel M, Pereira L, Riquelme A, Tolosana-Delgado R, van den Boogaart KG (2022) ***Inferring parameters of 3D particles microstructures from 2D sections using statistical learning***, in 21st Annual Conference of the International Association for Mathematical Geosciences – IAMG 2022, Nancy, France, Aug 29-Sep 3, 2022.

6. Avalos S, Ortiz JM, Leuangthong O (2022) ***Multivariate morphing transformation: Fundamentals and challenges***, in 21st Annual Conference of the International Association for Mathematical Geosciences – IAMG 2022, Nancy, France, Aug 29-Sep 3, 2022.

7. Riquelme AI, Ortiz JM (2022) ***A Riemannian tool for clustering of geo-spatial multivariate data***, in 14th International Conference on Geostatistics for Environmental Applications – geoENV 2022, Parma, Italy, June 22-24, 2022.

8. Larrondo P, Frank B, Ortiz J (2022) ***Automated topical extraction to aid in complex problem-solving feedback consistency in engineering design courses***. Proceedings of the Canadian Engineering Education Association (CEEA).

9. Koruk K, Ortiz JM (2022) ***Ensemble based domaining informed with unsupervised classification of geochemical data***, CIM 2022 Convention, Vancouver, BC, May 1-4, 2022.

10. Ortiz JM (2022) ***Predictive Geometallurgy and Geostatistics Lab***, invited talk, XUST-Queen's training program, December 1, 2022, online.

11. Ortiz JM (2022) ***Predictive Geometallurgy and Geostatistics Lab***, presented to Teck, November 2, 2022.

12. Ortiz J (2022) ***Geostatistics and geometallurgical Modelling***, in: Queen's Engineering Research Networking Day 2022, Queen's University, 12 October, 2022.

13. Ortiz JM (2022) ***Predictive Geometallurgy and Geostatistics Lab***, invited talk, Instituto de Ingenieros de Minas de Peru, March 14, 2022, online.

Conference posters

1. Tolosana-Delgado R, Avalos S, van den Boogaart KG, Frenzel M, Ortiz JM, Pereira L, Riquelme A (2022) ***Modelling microstructures with flexible Laguerre Mosaics***, in 21st Annual Conference of the International Association for Mathematical Geosciences – IAMG 2022, Nancy, France, Aug 29-Sep 3, 2022.

2. Riquelme AI, Ortiz JM (2022) ***Multivariate Simulation Using Locally Varying Corregionalization Models***, poster, CIM 2022 Convention, Vancouver, BC, May 1-4, 2022.

Funding

Research is possible thanks to the funding provided by Queen's University, NSERC through funding reference nos. RGPIN-2017-04200 and RGPAS-2017-507956, Mitacs Accelerate IT27769 in collaboration with SRK Consulting Canada, and MineSense support of Fouad Faraj work.

Predictive modelling workflows in geometallurgy[1]

Julian M. Ortiz (julian.ortiz@queensu.ca)

Abstract

Geometallurgy requires the integration of multiple data sources that are interrelated in complicated ways. Its aim is to provide models that link the spatial characterization of materials with process performance. These processes can be linked both to the mining extraction or ore and waste materials, as well as the processing of the ore, and even the disposal of the waste.

In this paper, a review of the methods and tools used to create predictive models in geometallurgy is provided. Different modelling workflows are presented and some case studies are included to illustrate how to tackle problems related to geological modelling, resource estimation, mine planning and mineral processing modelling. The workflows provide a map of how to face a modelling problem. The tools used at each step can be adapted to the type and quantity of data available, as well as the complexity of the problem. More importantly, these workflows can be integrated in a model of the entire system.

A review of the tools provided by statistical learning and geostatistics is included emphasizing the types of problems that can be tackled with each method. Also, the data requirements, challenges and limitations of the methods are described. Finally, these concepts are illustrated with practical applications, where the potential benefits obtained are highlighted as well as the assumptions and limitations.

1. Introduction

The sustainable management of mineral resources and reserves must maximize the benefits of the extraction of raw materials. These benefits involve economic revenues, as well as environmental and social aspects and ensure the continuous supply of metals and materials needed by society. Any mining operation combines diverse interconnected processes to go from in situ geological resources to final products that can be sold or used in the production of other materials and goods. In current operations, these processes are treated separately, managed independently, mostly predicted with deterministic models and updated in discrete time steps [Avalos, 2021]. This means that risks associated to the heterogeneity of these natural resources, as well as uncertainties linked to our incomplete knowledge of their properties and the associated processes, are not accounted for [Montiel et al., 2016]. The uncertainty is caused by the limited sampling information available to characterize the flows of materials in the mining system, as well as the complex nature of the physical and chemical processes these materials are subjected to recover the metals and materials of interest.

[1] Cite as: Ortiz JM (2022) Predictive modelling workflows in geometallurgy, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-01, 8-19.

This paper is an extended version of the paper "Predictive Modelling in Geometallurgy", presented in the 18th International Conference on Mineral Processing and Geometallurgy.

The sequential modelling of the processes in a mining system (resources, design, planning, drilling, blasting, loading, hauling, crushing, grinding, metallurgical processing) does not lend itself to their joint optimization [Ortiz et al., 2015]. In this paper, we present specific workflows for different stages of a mining system, and then discuss their possible integration and the advantages of this systemic modelling approach to geometallurgy.

2. Geometallurgy: an expanded definition

Geometallurgy aims at creating spatially-based predictive models by combining geological, mining and metallurgical information. These models are subjected to the mining, mineral processing and metallurgy processes and can be used anticipate their performance, control the processes and optimize the parameters and decisions involved [Ortiz et al., 2015]. Geometallurgy must be understood beyond process mineralogy, which has been the traditional context in which the term has been used [Dominy et al., 2018], and should account for the processes involved in the ore excavation.

Geometallurgical modelling aims at characterizing these processes in an interconnected framework that transfers properties of materials from one process to the next, accounting for their time of extraction, local (space and time) properties (grades, mineralogy, physical properties), blending that occurs in each process, residence time distribution, mass balance, etc. Each process can be modelled to obtain a prediction of the properties of the output, including estimated values and uncertainty quantification. Importantly, the transference of this uncertainty from one stage to the next should be modelled, accounting for correlations, interactions and blending. These models can have different levels of sophistication and complexity from simple statistical predictions to phenomenologic models.

A geometallurgical predictive model (which can be seen as a digital twin of the operation) can be created by combining realistic models of the different stages of the mining value chain, where particular inputs lead to outputs that feed other processes downstream. The resulting outputs can be calibrated with production information and be used to feedback into the models to enhance them [Benndorf and Jansen, 2017]. Optimization and control of the entire system is only possible if a complete understanding of all the components and their interactions is achieved.

3. Modelling workflows

In order to put together a geometallurgical predictive model, each process must be modelled as a stand-alone step, but all relevant inputs and outputs must be accounted for. Understanding which variables must be incorporated in each model is a result of specific domain knowledge.

Examples of modelling workflows can be found in the literature, linked to specific conditions or problems solved through modelling. These workflows will change depending on the type of data available and the specific conditions of the operation. However, typically mining projects need to deal with the following steps:

- Domaining
- Block model construction
- Mine planning and scheduling
- Processing and recovery

Although this is a high-level classification of the typical steps, the workflows can be adapted to specific constraints or requirements of the operation, including, for example, blending, economic constraints, water and energy management, environmental constraints, etc.

In the next sections, we review some possible workflows for each stage separately, and illustrate their modelling approaches, providing some real life illustrations of the results.

Domaining

The idea behind domaining is to identify subsets of samples with similar properties that can help create a spatially connected volume for a specific modelling purpose [Faraj and Ortiz, 2021]. Domains can be defined for resource, structural, or geometallurgical modelling, just to name a few. In each case, the relevant features change and the modeller must decide what aspects are relevant for the modelling goal, which is referred in **Error! Reference source not found.** as "domain knowledge". Multiple tools can be used in the clustering stage [Fouedjio et al., 2018]: constrained optimization, k-means, Gaussian mixture models, geostatistical clustering, etc. In real projects, this is often done manually, by checking which geological properties of the samples control the mineralization (or any other attribute of interest being modelled). This is an ill-posed problem, since in reality, the mineral resources are not divided into true domains, so there is no way to check that our definition of clusters (and later of domains) is appropriate, other than by a posterior validation.

For spatial modelling, once the samples have been labelled into one of K clusters, any categorical geostatistical estimation or simulation method can be used to define the extent of the domains. For instance, indicator approaches, truncated Gaussian or pluriGaussian simulation, or multiple-point simulation [Chiles and Delfiner, 2012], can be used to generate one or more domains models. If a single "expected" domain model is generated the approach is termed **deterministic**. If multiple possible scenarios are built that depict the possible extent of these volumes, then this is called a **stochastic** model.

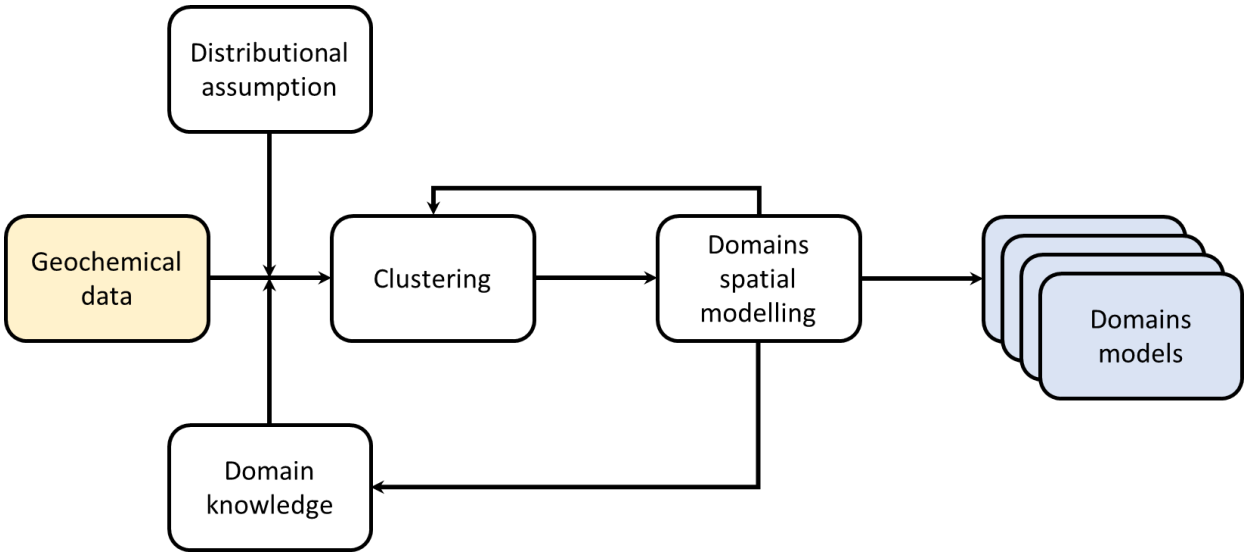


Figure 1: Example workflow for domaining.

An example of the results of an automated domaining approach is provided in Figure 2**Error! Reference source not found.**, where sample data are labelled according to their alteration domains, inferred from

geochemical data. These labelled samples are then used to determine the domains, with uncertainty, through an ensemble support vector classification [Koruk and Ortiz, 2022]. This approach combines domain knowledge (used in the determination of which geochemical variables are relevant to the model at hand), a simplified optimization technique to assign samples to each cluster, controlled by a match to the parametric distribution of these variables in each domain, and a machine learning method for classification, which in this case is extended to provide multiple scenarios of the extent of the volumes classified in each domain. The output of this workflow, can be a stochastic model, represented by an ensemble of models of the domains, or as a single probabilistic model, i.e. where for each block, the probability of belonging to each category is quantified, or a deterministic model, represented by the most likely category in each block, determined through majority voting over the ensemble of models, or some other consensus.

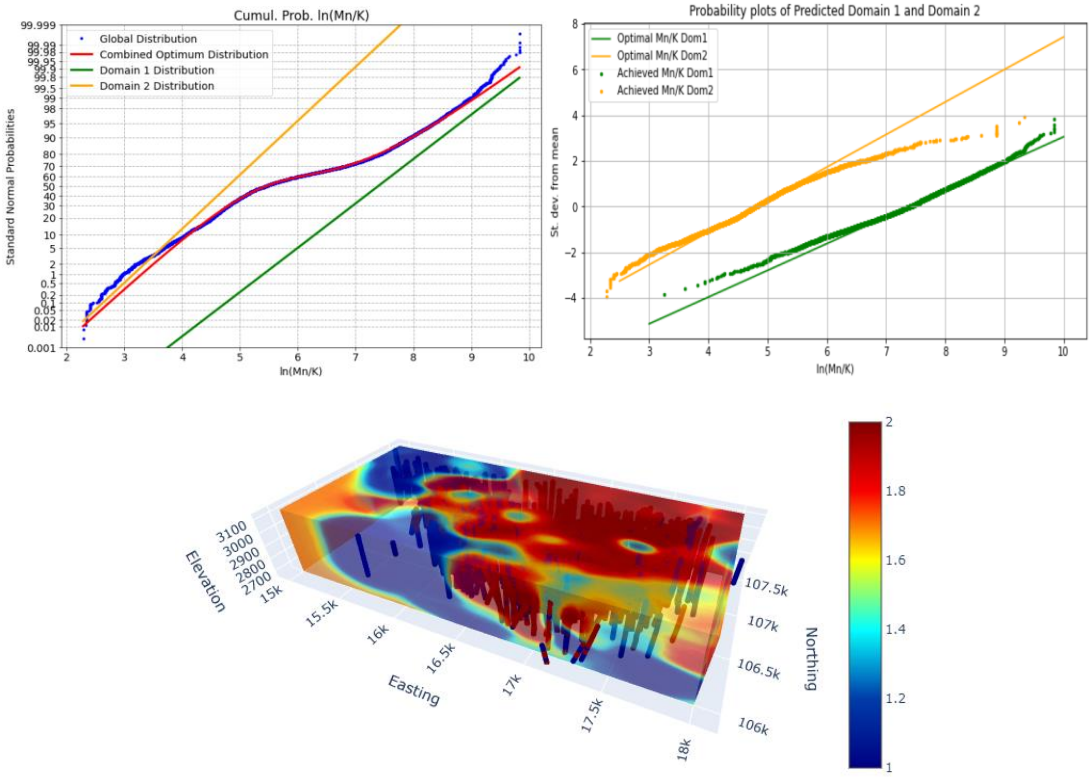


Figure 2: Samples are assigned to categories: distributional assumption (top left), clustering (top right). A probabilistic model is built with SVC using resampling and bagging (bottom).

Block model construction

Once the domains have been established, the construction of a spatial block model follows, by estimating or simulating the different attributes within each domain volume, constrained by the sample information belonging to that same domain. This is conventionally done with geostatistical tools such as kriging or conditional simulation. Accounting for multivariate relationships is key to capture interactions that are relevant for process modelling downstream. These multivariate relationships are both statistical, i.e. correlations between variables collocated, and spatial, measured through direct and cross-variograms between the different variables within each domain.

There are well established techniques for multivariate modelling, including geometallurgical attributes [Deutsch et al., 2015], compositional data (such as elemental or mineral proportions) [Tolosana-Delgado et al., 2019] or multiple variables with complex relationships [Barnett et al., 2014; Avalos et al., 2022].

Workflows for this stage are well-established and thoroughly documented in the literature. As an example, a typical workflow using a transformation into factors that can be modelled independently in space, is provided in Figure 3**Error! Reference source not found.**. The input considers multiple domain models (that account for uncertainty in the extent of the domains) and multivariate sample data representing grades. The result is a set of block models with the spatial distribution of the grade models, respecting their complex relationships, and their spatial continuity, and honouring the data.

A bivariate example of this is illustrated in Figure 4, for two synthetic variables with a complex non-linear correlation. The original variables are transformed to "morphing factors", which are spatially uncorrelated and can be simulated independently with any geostatistical simulation method. The back-mapping provides realizations that honour the data, and spatial and statistical relationship between the two attributes [Avalos et al., 2022]. This technique has been tested with up to 10 variables, providing excellent reproduction of the statistical and spatial correlations.

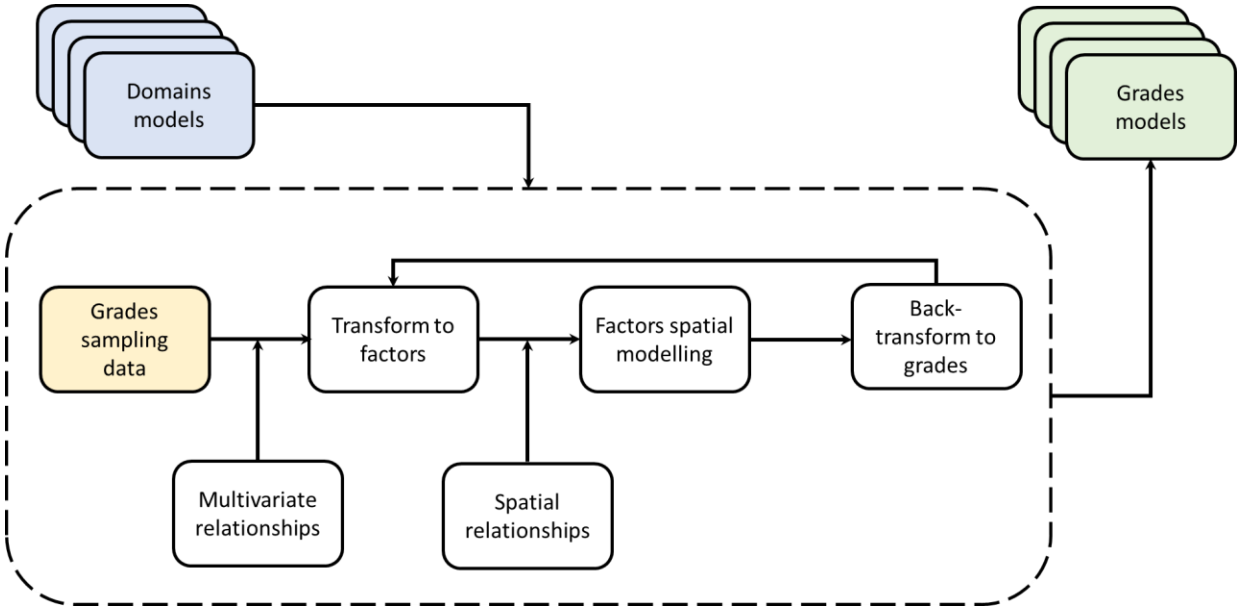


Figure 3: Example workflow for block model construction.

As with the domaining approach, the block model construction workflow allows for the generation of a deterministic or a stochastic output. The deterministic model, which is similar to an estimation, is obtained by locally averaging the realizations obtained through simulation. This is known as an E-type model. Alternatively, the realizations can be used to determine a local probability distribution at every location, or the full ensemble of realizations can be kept to input in subsequent steps.

Finally, it is important to note how stochastic input domain models are used. Rather than propagating uncertainty by using each domain input model and creating an ensemble of realizations of grade distributions within these domains, a cascade approach is taken, where a single realization of the grades is generated for each domain model. In this fashion, if we have 100 domain models at the beginning, the

product of the block model construction stage is 100 models of grade distributions, each constrained by a different domain model. This allows us to sample the space of uncertainty with a limited number of models, avoiding the explosion in the number of models if these are nested.

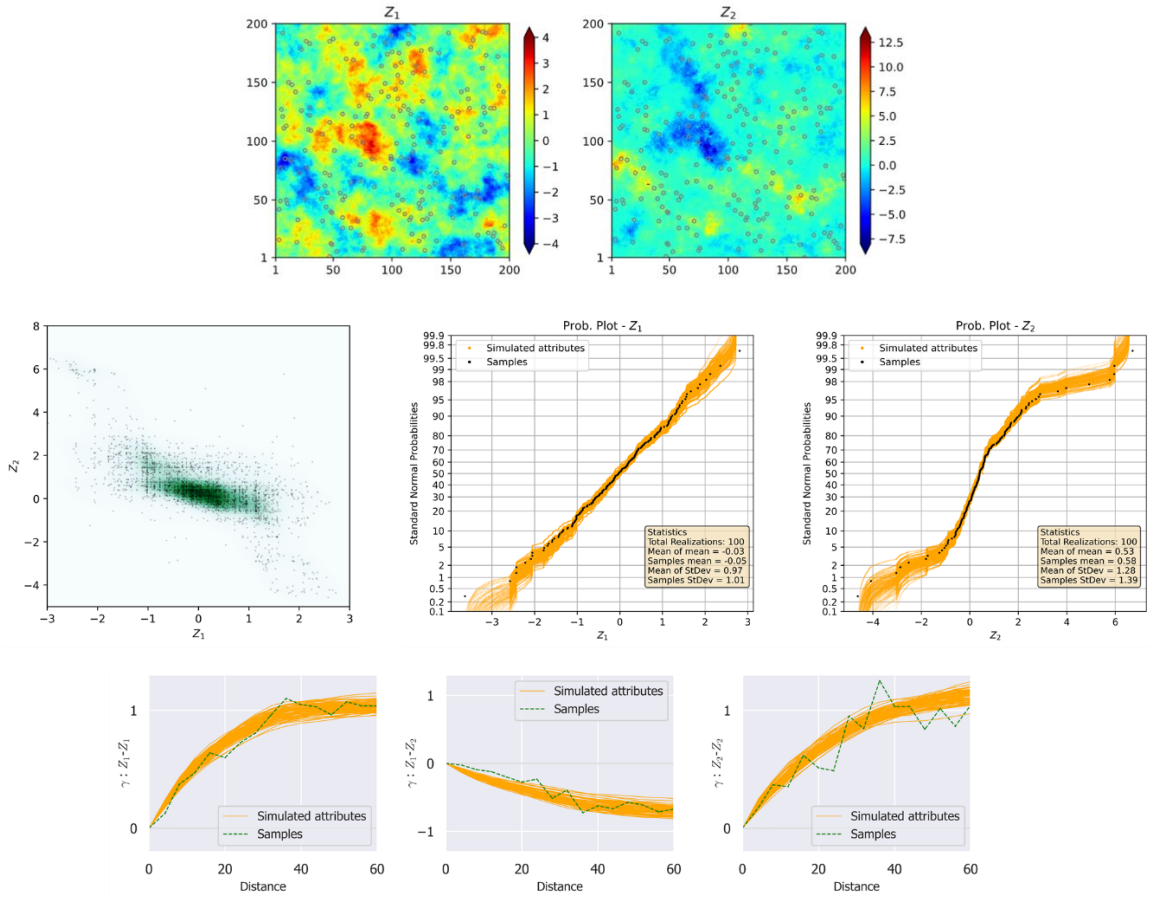


Figure 4: Multivariate block model construction.

Mine planning and scheduling

Mine planning and scheduling define the decision of extracting blocks and the sequence over time and space of this process. Planning must account for the uncertainty in the spatial distribution of the attributes of interest and multiple constraints related to their extraction (economic, operational, environmental, geometric, etc.). There are multiple approaches to optimizing the mine plan, including stochastic optimization [Dimitrakopoulos, 2011]. Scheduling can be done with stochastic integer programming accounting for uncertainty of the attributes [Morales et al., 2019], and constraints can be inputted in the model (see for example an approach to account for demand-side management for energy consumption [Diaz et al., 2016]). Some of the newer proposals involve reinforcement learning approaches to train the system optimizing sequential decision making [Avalos and Ortiz, 2021]. A workflow for a trained Deep Q-Learning neural network is shown in Figure 5. A set of grades models are combined to provide a grades model with uncertainty, which along with production data is used to train a Deep Q-Network. The grades models, in this particular workflow, are processed into an E-type model and a model of the local variance, thus representing the block uncertainty distribution by these two parameters. The network then learns

the probability of extraction for each feasible block. Based on these probabilities, a decision about the next block to extract is made, and this is used to assess the quality of the decision and update the probabilities for the remaining blocks. The neural network evolves from a stage where different trials are applied (called exploration phase) so that it can learn what the effects of new decisions are, and slowly moves towards an exploitation phase, where mostly decisions whose outcomes are known, are applied. In this fashion, the network learns to make the best decisions and provides an optimized extraction sequence. Interestingly, this approach can also incorporate feedback from production, thus allowing a learning strategy that rewards decisions that in practice have a positive impact in the objective function.

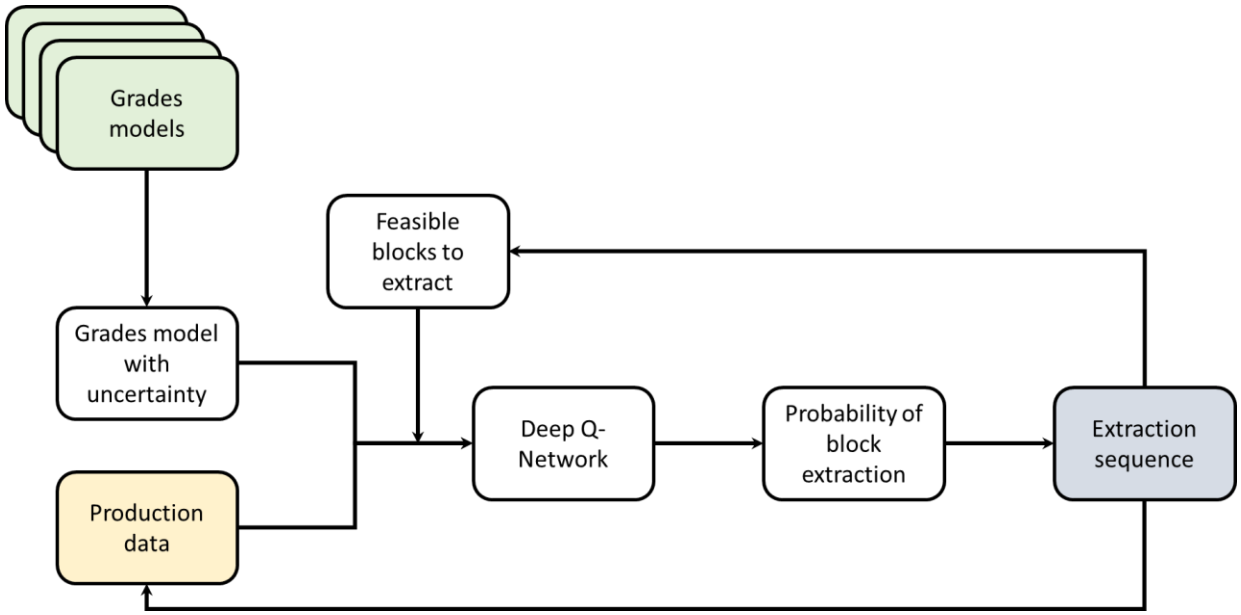


Figure 5: Example workflow for mine scheduling.

The results of the methodology are illustrated in Figure 6, where a schedule is defined randomly initially and after the neural network learns the probabilities of extraction, based on its training, decides a better schedule that increases value in about 5% [Avalos and Ortiz, 2021].

At this stage, the uncertainty from domaining and from the grades distribution is accounted for. The optimization of the plan, considers these two sources of uncertainty and comes up with a single strategy tailored to manage the possibilities given by the available information. The resulting plan can then be applied to each one of the grades models, in order to assess the impact of the domains and grades uncertainty in the production outcome from the mine. These time series of blocks extracted in a fixed sequence, for each realization of the grades distribution, can be used as the processing streams inputted into the crusher and the processing plant.

Processing and recovery

As mentioned, the sequence of extraction and schedule from the mine plan provide the transfer function to go from the spatial domain to the time domain. The blocks extracted from the mine become truckloads fed to the processing streams. Transferring the uncertainty can be done by use of simulation or considering a multiGaussian kriging approach [Riquelme and Ortiz, 2021]. Once the time series characterizing the materials in each processing stream are available, the processing performance can be

predicted. Since there are many different processing flowsheets, models will vary. Examples of such models involve hardness prediction with deep learning [Avalos et al, 2020], AI driven air classification of particles [Otwinowski et al., 2022], deep learning froth flotation recovery prediction [Pu et al., 2020], and recovery prediction in leaching using machine learning [Flores and Leiva, 2021].

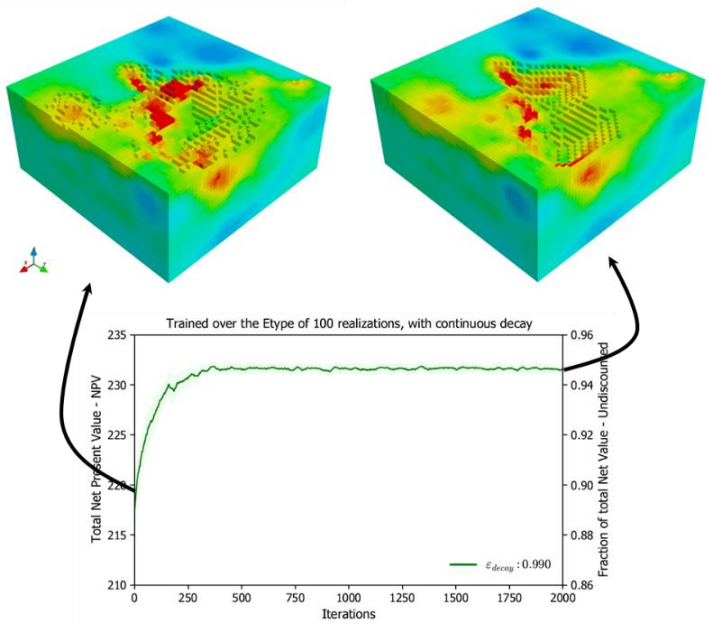


Figure 6: Illustration of the schedule before and after training.

An interesting compilation of models in mineral processing can be put together as a discrete event simulation [Moraga et al., 2022]. For each grades sequence coming from the expected schedule, the changes at each stage of the process can be modelled, in order to predict residence time distribution, particle size distribution and properties of the streams of concentrate and tailings in flotation. This is illustrated in Figure 7, where the processing stages are interconnected in a sequence, producing two outputs: concentrate and tailings. The models for crushing and grinding will vary depending on the equipment, type of circuit configuration and settings used. Similarly, flotation involves recirculation of streams, which have an effect on the resulting products. This can be modelled as well, for a particular configuration, as depicted in Figure 8.

Modelling the processing stage allows for the prediction of blends that occur during grinding and flotation, and are an important input if the processing stage is to be controlled for optimum output.

Integrated model

From the workflows presented earlier, it is possible to create an integrated model from the original drilling data, to the processing stage, considering the concentrate as a product, and also allowing the management of the tailings. This is illustrated in Figure 9.

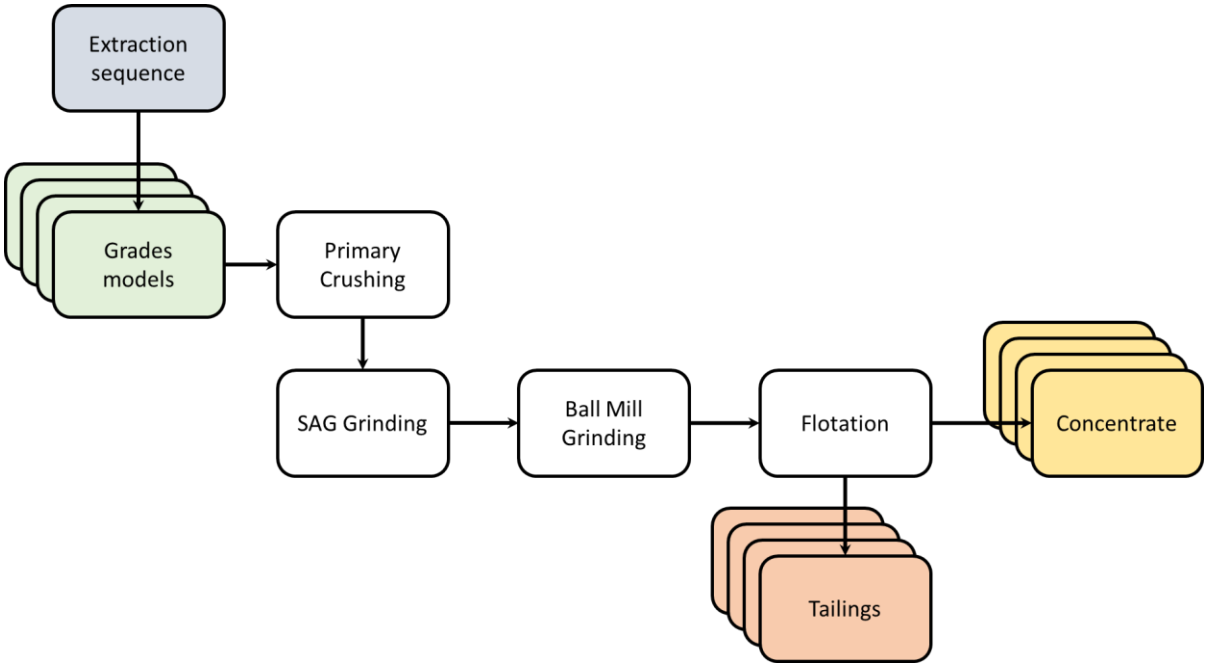


Figure 7: workflow for modelling performance during processing.

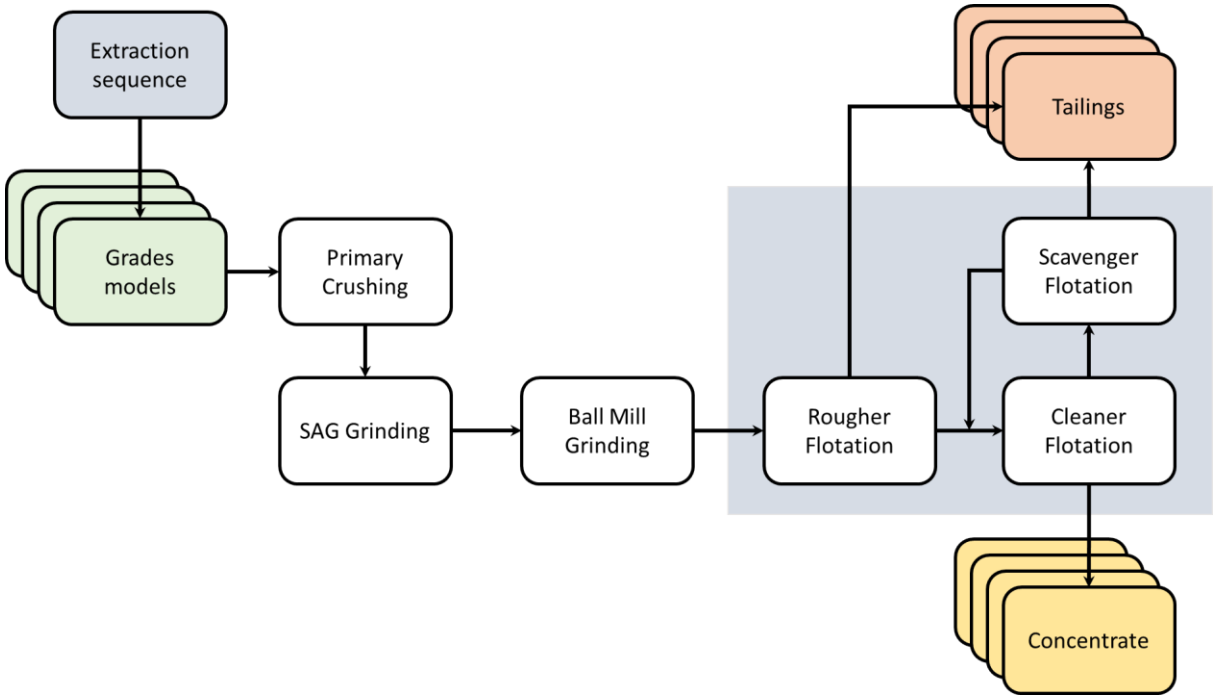


Figure 8: workflow for modelling performance during processing, for a particular flotation circuit configuration.

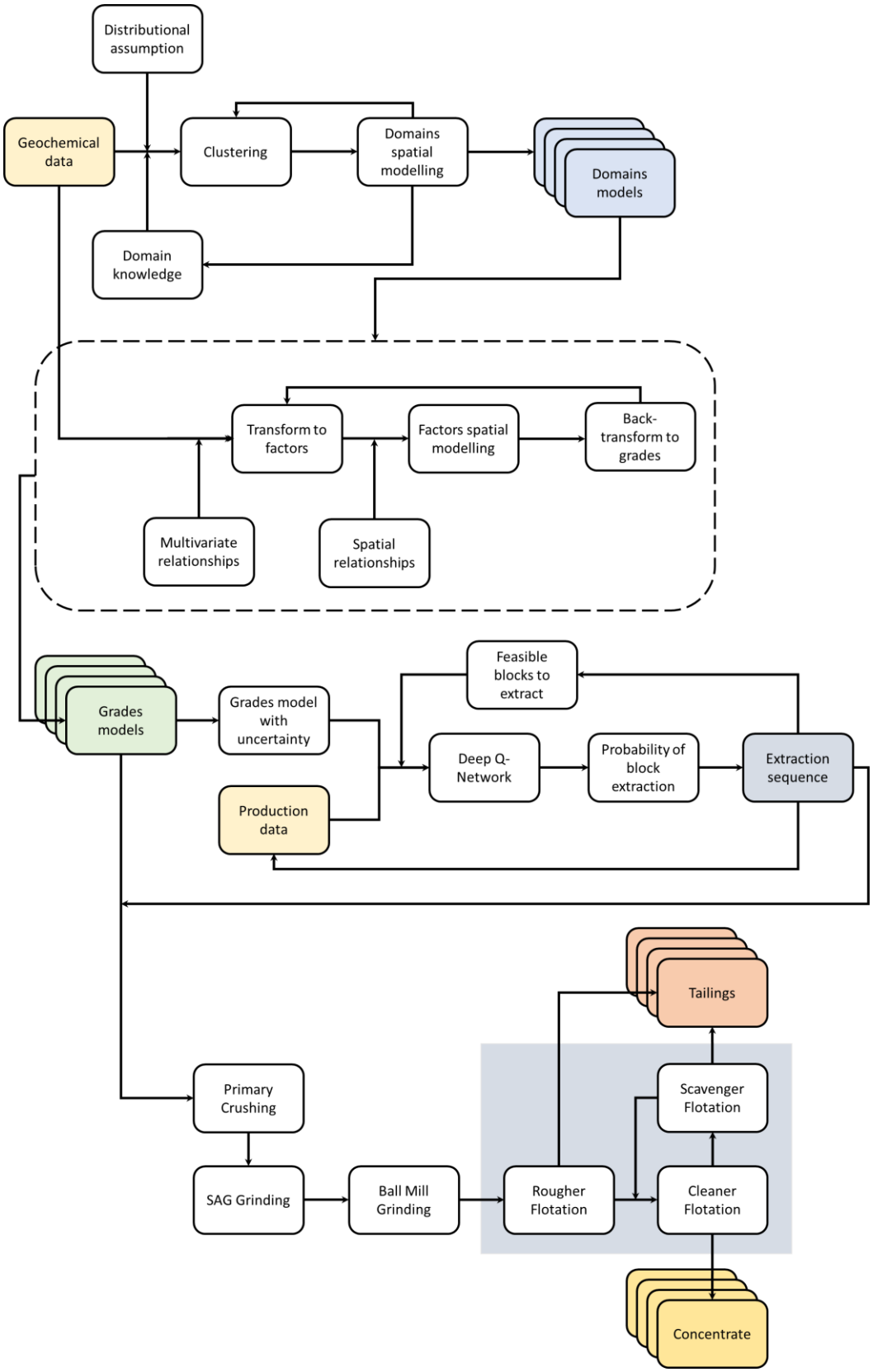


Figure 9: Workflow of integrated process, from drillholes to concentrate.

4. Conclusions

A mining system can be seen as a sequence of stages or processes. The actual ore deposit, mine and operation is modelled through the block model for resources, reserves and the mine extraction schedule. Mineral processing and the metallurgical process can also be modelled, accounting for the variability and uncertainty in the feed and in the corresponding process. All this information can be used to perform optimization and control over each stage of the process, but, most importantly, once the geometallurgical model is fully implemented and connected, the model can be optimized globally and multiple constraints can be incorporated in the decision making process.

Given the complexity of stages and processes in mining, the behavior of each component of the system can only be approximated. Real time sensors, composited measurements, and soft sensors can be used to update the status of the model. This predicted output must be compared to other measurements of the actual output, for the models to learn. Feedback must be constant to keep models updated at a time scale relevant to decision making. Prediction should incorporate uncertainty as a key factor, to ensure the properties of the final product and the waste are anticipated. This approach ensures all considerations are accounted for when extracting raw materials and recovering the elements of interest, thus making this process more sustainable.

5. References

Avalos S (2021) Advanced predictive methods applied to geometallurgical modelling, Ph.D. thesis, Queen's University.

Avalos SA, Kracht W, Ortiz JM (2020) An LSTM approach to SAG mill operational relative-hardness prediction, Minerals, 10 (9), 734. https://doi.org/10.3390/min10090734

Avalos S, Ortiz JM (2021) Geometallurgical modelling and deep Q-Learning to optimize mining decisions, in 11th International Geostatistics Congress, Toronto, Canada.

Avalos S, Ortiz JM, Leuangthong O (2022) Multivariate morphing transformation: Fundamentals and challenges, in 21st Annual Conference IAMG 2022, Nancy, France.

Barnett RM, Manchuk JG, Deutsch CV (2014) Projection pursuit multivariate transform. Mathematical Geosciences, 46 (2), 337–359. https://doi.org/10.1007/s11004-013-9497-7

Benndorf J, Jansen JD (2017) Recent Developments in Closed-Loop Approaches for Real-Time Mining and Petroleum Extraction, Mathematical Geosciences, 49, 277-306. https://doi.org/10.1007/s11004-016-9664-8

Chiles J-P, Delfiner P (2012) Geostatistics – Modeling Spatial Uncertainty, Second Edition. Willey, 699 p.

Deutsch JL, Palmer K, Deutsch CV, Szymanski J, Etsell TH (2015) Spatial Modeling of Geometallurgical Properties: Techniques and a Case Study, Natural Resources Research, 25(2), 161-181.

Diaz S, Gil E, Palacios A (2016) Demand side management, an imperative asset for the mining industry, ENERMIN 2016: 3rd International Seminar on Energy Management in Mining, Santiago de Chile, 2016, p 10.

Dimitrakopoulos R (2011) Stochastic optimization for strategic mine planning: a decade of developments. Journal of Mining Science 84, 138–150. https://doi.org/10.1134/S1062739147020018

Dominy SC, O'Connor L, Parbhakar-Fox A, Glass HJ, Purevgerel S (2018) Geometallurgy—A Route to More Resilient Mine Operations, Minerals 8(12): 560. https://doi.org/10.3390/min8120560

Faraj F, Ortiz JM (2021) A simple unsupervised classification workflow for defining geological domains using multivariate data, Mining, Metallurgy & Exploration, 38, 1609-1623.

Flores V, Leiva C (2021) A Comparative Study on Supervised Machine Learning Algorithms for Copper Recovery Quality Prediction in a Leaching Process. Sensors, 21(6): 2119. https://doi.org/10.3390/s21062119

Fouedjio F, Hill EJ, Laukamp C (2018) Geostatistical clustering as an aid for ore body domaining: Case study at the Rocklea Dome channel iron ore deposit, Western Australia. Applied Earth Science, 127(1), 15–29. https://doi.org/10.1080/03717453.2017.1415114

Koruk K, Ortiz JM (2022) Definition of geological domains with Ensemble Support Vector Classification, in 21st Annual Conference IAMG 2022, Nancy, France.

Montiel L, Dimitrakopoulos R, Kawahata K (2016) Globally optimising open-pit and underground mining operations under geological uncertainty. Mining Technology, 125, 2–14. https://doi.org/10.1179/1743286315Y.0000000027

Morales N, Seguel S, Cáceres A, Jélvez E, Alarcón M (2019) Incorporation of Geometallurgical Attributes and Geological Uncertainty into Long-Term Open-Pit Mine Planning. Minerals, 9(2): 108 https://doi.org/10.3390/min9020108

Ortiz JM, Kracht W, Townley B, Lois P, Cárdenas E, Miranda R, Álvarez M (2015) Workflows in Geometallurgical Prediction: Challenges and Outlook, in 17th Annual Conference IAMG 2015, Schaeben H, Tolosana-Delgado R, van den Boogaart KG, van den Boogaart R (Eds.), 6 p.

Otwinowski H, Krzywanski J, Urbaniak D, Wylecial T, Sosnowski M (2022) Comprehensive Knowledge-Driven AI System for Air Classification Process. Materials, 15(1): 45. https://doi.org/10.3390/ma15010045

Pu Y, Szmigiel A, Chen J, Apel DB (2020) FlotationNet: A hierarchical deep learning network for froth flotation recovery prediction. Powder Technology, 375, 317–326. https://doi.org/10.1016/j.powtec.2020.07.102

Riquelme AI, Ortiz JM (2021) Uncertainty assessment over any volume without simulation: revisiting multi-Gaussian kriging, Mathematical Geosciences, 53, 1375-1405. https://doi.org/10.1007/s11004-020-09907-9

Tolosana-Delgado R, Mueller U, van den Boogaart KG (2019) Geostatistics for Compositional Data: An Overview. Mathematical Geosciences 51, 485–526. https://doi.org/10.1007/s11004-018-9769-3

Testing a new sequential isofactorial simulation algorithm[1]

David Casson (3drc1@queensu.ca)
Julian M. Ortiz (julian.ortiz@queensu.ca)

Abstract

Existing geostatistical simulation approaches frequently rely on an assumption of Gaussianity or a computationally intense interpolation method to produce simulation realizations. These simulation methods encounter challenges in reflecting the attributes of sample data set exhibiting skewed distribution or non diffusive (clustered) properties. A new simulation algorithm has been developed, following the Sequential Isofactorial concept identified by Emery (2002). The simulation algorithm, referred to as the Sequential Isofactorial Algorithm ("SIA") allows for choice of a multivariate Gaussian or multivariate gamma based model to reflect a symmetrical or skewed distribution respectively. The SIA also incorporates a choice of destructuration coefficient reflecting varying degrees of clustering in the data (non diffusiveness in the multivariate distribution). Initial results from conditional simulation with the SIA utilizing a handful of conditioning data points suggest the results reflect the attributes of the chosen model, exhibiting the appropriate histogram, variogram and clustering. A comparison to sequential Gaussian simulation ("SGS") also creates an intuitive result, with the new algorithm showing greater variance in realizations. The SGS algorithm is constrained to produce a Gaussian local conditional cumulative distribution function with mean based on local conditioning data and variance independent on the sample values, only dependent on the spatial configuration of the information. The SIA (under a Gaussian diffusive model) also produces a Gaussian local CCDF with mean based on local conditioning data, however, the variance is dependent on the variance of the local conditioning data values.

1. Introduction

1.1. Mineral Resource Estimation and Traditional Simulation Techniques

Traditionally, the qualified person completing a Mineral Resource estimate applied kriging to arrive at a best estimate for mineral resource grade and tonnage. Kriging is a deterministic method unable to provide a relevant measure of uncertainty associated with its deterministic estimates. The qualified person would classify the mineral resource as measured, indicated, or inferred based on personal experience and industry association guidance related to drill hole spacing and deposit type.

Simulation provides a quantification of uncertainty in resource modelling. This allows an estimator to arrive at a quantitative view of how probable it is that a given block has a given value (e.g., probability that grade of block is above cut off). This probabilistic estimate of uncertainty can be carried through the

[1] Cite as: Casson D., Ortiz J. M. (2022) Testing a new sequential isofactorial simulation algorithm, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-02, 20-46.

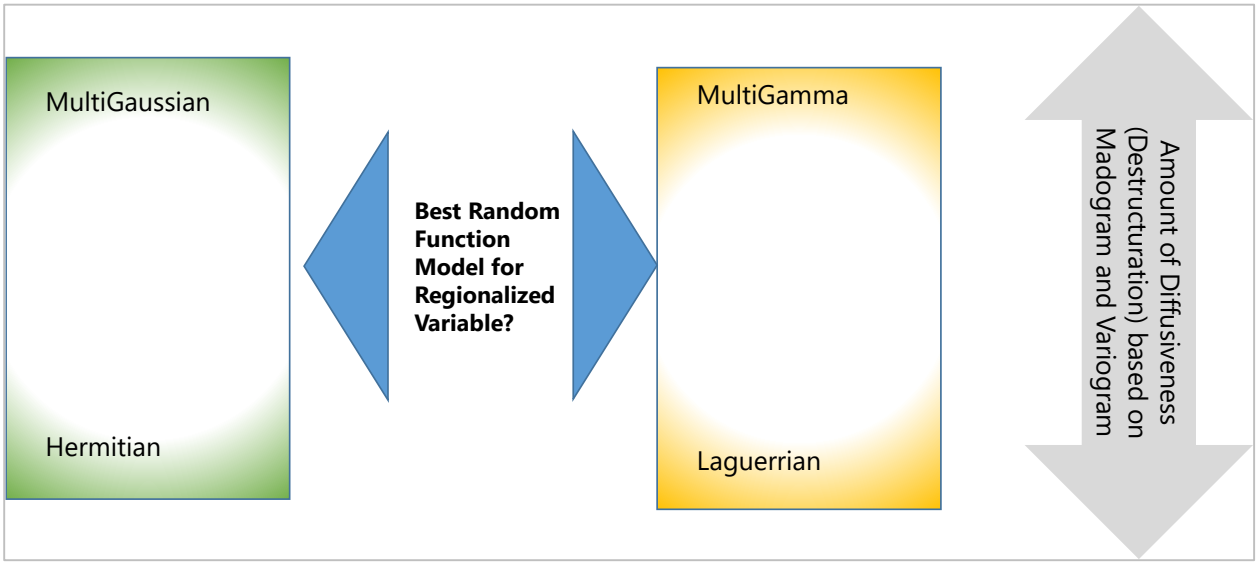
mine design and operation to make better decisions, understanding the range of possible operational and financial outcomes and associated likelihood.

Commonly applied simulation techniques include Sequential Gaussian Simulation and Sequential Indicator Simulation. Both these sequential simulation methods require a random draw from an estimated Conditional Cumulative Distribution Function ("CCDF") at unsampled locations. The CCDF, which is based on information in a neighborhood deemed to be relevant, reflects the probability that the unknown value at the location in question is below any value in its possible range. Simulation will not be successful if the CCDF is not accurate. Sequential Gaussian Simulation relies on the assumption that the regionalized variable under study is multiGaussian in nature. This is often not true. Simulation will also not be successful if it is not practical to implement. Sequential Indicator Simulation makes no assumption on the model of the variable under study but requires computation and modelling of many indicator variograms which makes it a cumbersome approach and challenging to implement in practice. Sequential Indicator Simulation also fails to account for cross correlation of indicator values at various thresholds, resulting in a sub-optimal model for the local CCDF (Emery and Ortiz, 2004; Machuca-Mory et al., 2008).

1.2. Context on Developed Simulation Isofactorial Algorithm

This broader research effort proposes to create a practical tool to choose either a multivariate Gamma or multivariate Gaussian random function model and associated parameters including level of destructuration based on observed properties of the available sample data. The "informed selection" of random function model, as shown in Figure 1 below can be applied in simulation of ore grades through the SIA. Importantly, this work will empirically test the relative success of this approach on partially sampled but known exhaustive data sets. Results will be compared against the exhaustive data set as well as simulation results obtained when traditional sequential Gaussian simulation is applied to the same sampled data set.

Figure 1: Illustration of Random Function Model Parameter Selection



1.3. Disjunctive Kriging to Model Gaussian or Gamma Local CCDF Functions

The Sequential Isofactorial Algorithm ("SIA") relies on disjunctive kriging based sequential simulation utilizing asymptotic polynomial expansions to model local conditional cumulative distribution functions ("CCDF's"). The polynomial expansion values calculated for sample values/locations can be used in conjunction with simple kriging to determine the transformed (e.g., normal score or "gamma score") value at an unsampled location. Disjunctive kriging is equivalent to simple co-kriging of the polynomial expansion terms. The specific polynomial expansions considered are orthogonal basis and the covariance between polynomials of different orders is zero. As a result, disjunctive kriging becomes simple kriging of the polynomial values for the transformed data values and then a linear weighted sum of the resulting polynomial values across orders. Disjunctive kriging is the equivalent of full indicator co-kriging but avoids the challenges of developing a workable model of co-regionalization for the indicator thresholds. This approach allows for the calculation of an expected value at an unsampled location based on statistical distance of nearby data points and reflecting the local conditioning data itself as embedded in the polynomial expansion terms at nearby sampled locations. The polynomial approximation technique can be used to define an expression for the local CCDF. Polynomial expansions may reflect a choice of multivariate random function family. The approach can be applied to sequentially simulate non multiGaussian data sets, such as multi-gamma data sets.

1.4. Randomizing Correlation Coefficient to Reflect Clustering of Values

Emery (2008) identified that the destructuration of grade (i.e., non diffusiveness of the random function model) can be reflected in simulation realizations by "randomizing" the correlation coefficient used in the disjunctive kriging process at each order of the polynomial expansion. Chiles and Delfiner (2012) provide a good overview of this technique. This approach takes the pure Gaussian and Gamma model and extends them to the more generalized "Hermitian" and "Laguerrian" models respectively. Without this adjustment, the Gaussian and Gamma models would be described as pure diffusive models. This diffusive property requires that the correlogram of the polynomials of order "p" are equal to the correlogram of the variable raised to the power p. Practically, the diffusive property means that sharp transitions (e.g., connectivity of extreme values) cannot be reflected in the model. In a diffusive model the transition between data points of different values will be gradual vs. abrupt. This is due to the higher order correlograms trending to zero as the power p increases (pure nugget). The incorporation of destructuration allows higher order terms to have greater weighting and, as a result, creates potential for sharper (non-diffuse) transitions, that appear in data sets as clustering of values.

2. Simulation Algorithm

2.1. Isofactorial Simulation

Polynomial expansion based disjunctive kriging can be used to define an expression for the local CCDF. This allows for the local CCDF to reflect the choice of bivariate random function family. Ortiz (2004) provides a good overview of fitting a finite function (any finite function is acceptable) with an expansion of Hermite Polynomials. A similar approach can be used to fit a finite function with an expansion of Laguerrian Polynomials.

The polynomials up to a selected order "N" are calculated at all sampled locations based on normal score or gamma score values. A variogram model is fitted to values. The covariance between polynomials of an

order N is based on the variogram with the resulting correlogram value raised to the power "N". Simple kriging is completed to solve for the numerical Hermitian or Laguerrian polynomial values of each order at the unsampled location.

For a given unsampled location we now have a set of numerical values for the polynomials at each order N. These polynomial values form the building blocks for the local CCDF at the unsampled location, effectively "coding" the information contained in the conditioning values (values and variance). The local CCDF is a weighted sum of these polynomial values. The weightings for each order polynomial in the CCDF are derived based on an expansion of the indicator function (i.e., probability value is less than or equal to a given threshold). The result is an equation for the CCDF value at an unsampled location as a function of the actual (but unknown) random variable value at that location.

2.2. Randomizing the Correlation Coefficient

The above procedure for disjunctive kriging considers only the case of no destructuration (i.e., diffusive models) as the correlograms of various orders are simply the correlogram raised to the power "N". This process can be adjusted to consider non-diffusive models. This is accomplished by a randomization of correlogram at higher orders using a beta distribution whose form is dictated by a specific scalar factor. As the scalar factor varies between zero and infinite, the model varies between the diffusive and non-diffusive (mosaic) case. In effect, as the model moves away from a pure diffusive model, greater weight is given to higher order expansions, resulting in a "tighter" local CCDF with less variance around its mean (i.e., sampling form the CCDF is more likely to yield a value closer to the local conditional mean and less informed by the broader background global mean). The incorporation of destructuration is accomplished by this change to the correlogram used in simple kriging of the polynomial values. The resulting numeric polynomial values kriged for a given unsampled location reflect both the underlying choice of bivariate family (i.e., Gaussian or Gamma) and a chosen amount of destructuration (clustering).

2.3. Sequential Simulation Algorithm

As described above, we can use the Isofactorial disjunctive kriging approach to create an equation for the local CCDF value as a function of the unknown underlying value. Sequential simulation draws a random number between zero and one and assumes this to be the CCDF value. Because the CCDF (by definition) is a monotonic function, a guess and check bounding algorithm was designed to iteratively determine the corresponding actual variable value that the CCDF random draw corresponds to (within a specified tolerance). The algorithm for solving the underlying value that corresponds to the randomly drawn CCDF value is described below.

1. Complete a random draw between one and zero for the simulated CCDF percentile value ("simulated percentile")
2. Based on range of known values (i.e., normal or gamma score values at sample locations) select a maximum and minimum possible value
3. For each of the maximum value, minimum value and midpoint value calculate the CCDF percentile using the modeled local CCDF function
4. Compare the simulated percentile value to the percentile value calculated at maximum, minimum and midpoint of the underlying normal or gamma scores

5. If simulated percentile value is within an acceptable tolerance of one of these percentile values, set the associated underlying normal score value as the simulated normal or gamma score value at that location, otherwise:
    a. Determine if the simulated CCDF value lies above or below the midpoint CCDF value, if above, set the minimum underlying value to the midpoint value, if below set the maximum underlying value to the midpoint value
    b. Go back to step 3 using this new minimum, maximum and midpoint value for the underlying
    c. Repeat until simulated percentile value is within acceptable tolerance of maximum, minimum or midpoint CCDF values and when it is set the simulated underlying value to the corresponding minimum, maximum or midpoint underlying value

Once the above algorithm is complete, we move (randomly) to the next unsampled location and repeat the algorithm considering both sampled locations and previously simulated locations in the neighborhood of the location being simulated.

3. Results

3.1. Examining Local CCDF Curves

A small five by three matrix was populated by randomly drawn values with three "unsampled" locations left blank in the array. This simple data set was used to check that modelled local CCDF curves were behaving in a manner that was intuitively expected. The results are shown below in figures 2, 3, 4 and 5.

Figure 2: CCDF Curves at Unsampled Points Demonstrating Impact of Conditioning Data

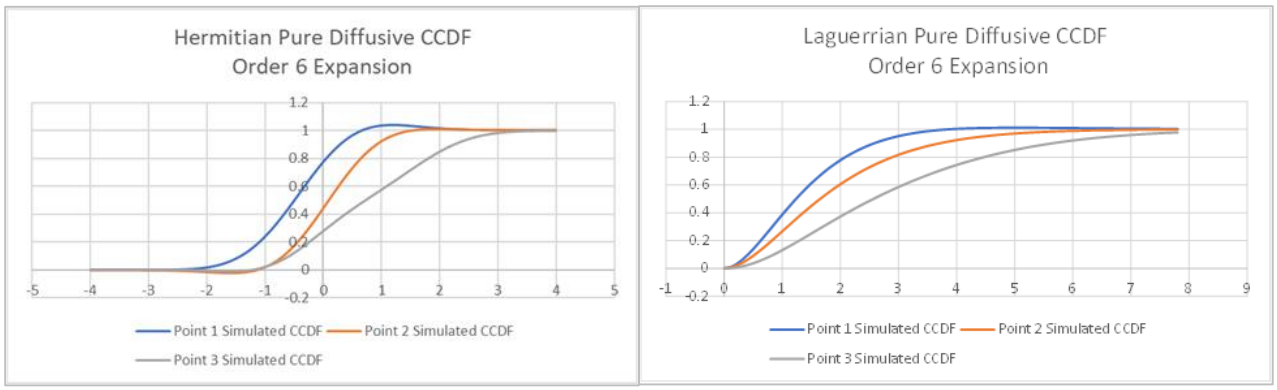


Figure 3: Illustration of CCDF Curves under Laguerrian cases with varying shape assumption.

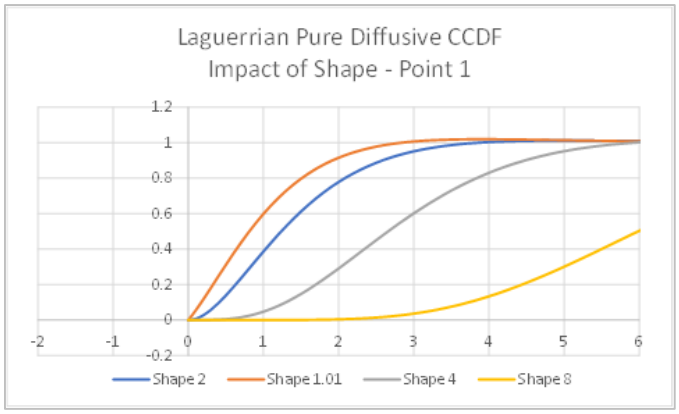


Figure 4: Illustration of CCDF Curves under Hermitian case with varying order of expansion

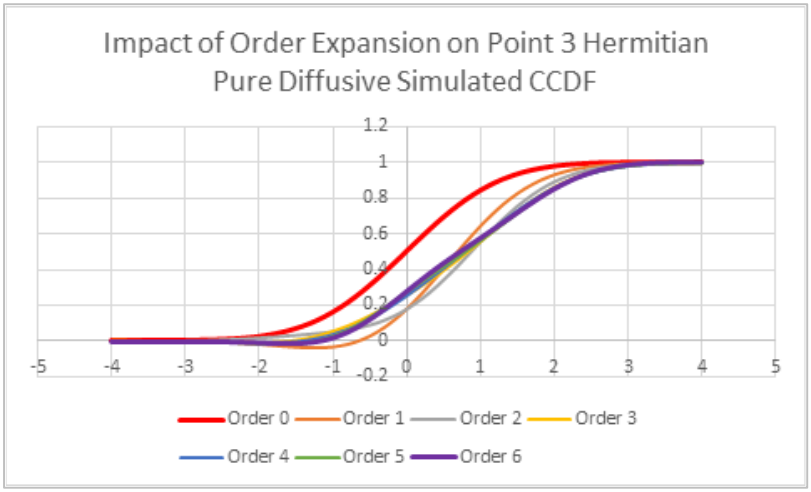
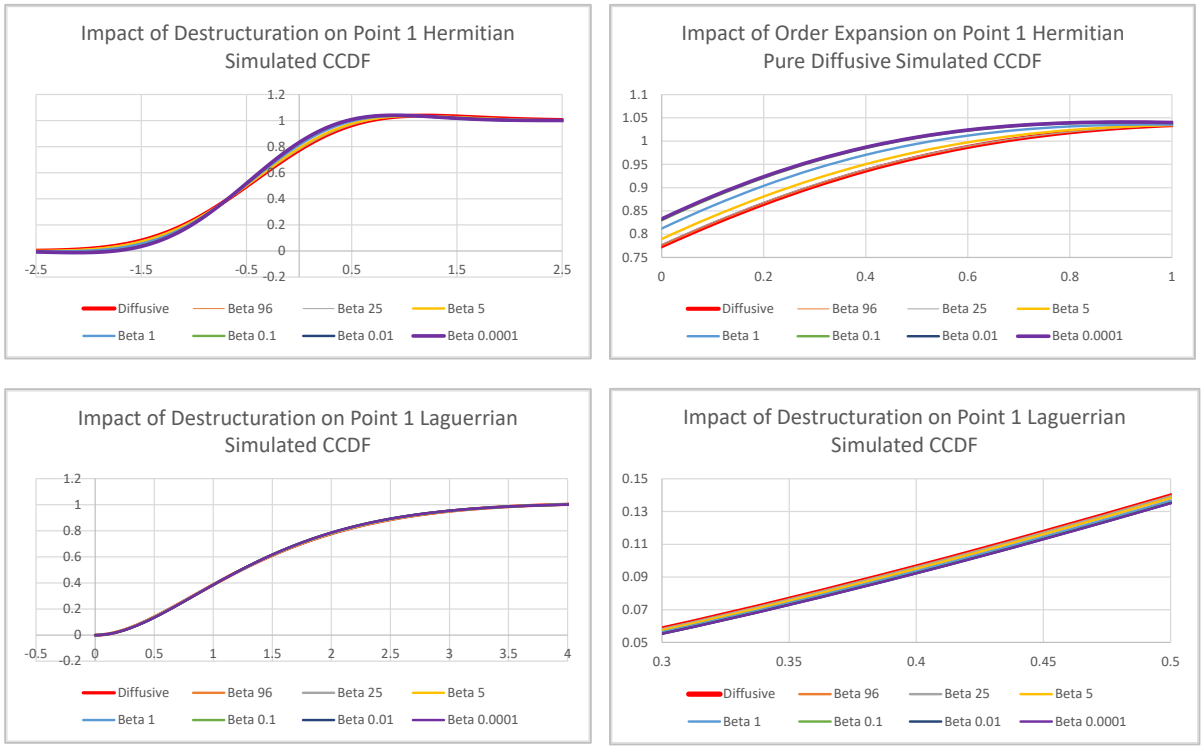


Figure 5: Illustration of CCDF curves at 1st simulated node under Hermitian and Laguerrian cases with varying destructuration



3.2. Comparison of SIA to Sequential Gaussian Simulation

Fifty simulation realizations on a 50x50 grid using the same conditioning data point were run with both Sequential Isofactorial Algorithm ("SIA") (Gaussian pure diffusive model) and traditional gslib SGSIM ("SGS").

- Single conditioning data point set at center of matrix of 0.2
- Lognormal reference distribution used resulting in normal score conditioning data point of ~-1.6 (based on lognormal percentile of 0.2 value)
- Spherical variogram with range of 8, nugget of 0.1
- Search radius of 12

While SIA results generally match those of SGS, SIA displays considerably greater variance than SGS, evidenced in variogram sills, conditional variance map and simulation realizations. This is driven by the fact that SGSIM is limited to a Gaussian CDF at any node, with variance obtained by simple kriging, hence independent of the sample values. In the SIA algorithm, the local CCDF's are built from the polynomial expansion. It is observed that the first term in the isofactorial expansion is the standard distribution, the second term adjusts for local mean, and subsequent terms adjust for the specific characteristics of the local conditioning data (e.g., a high variance in local conditioning data is reflected in the local CCDF curve being lower slope). The resulting variance of the CCDF depends on the particular sample values. Simulation results (in transformed normal score) are shown below in figure 6, 7, 8, 9 with SGSIM results on the left and SIA results on the right.

Figure 6: Comparison of SGS and SIA Realizations and Average (E-Type)

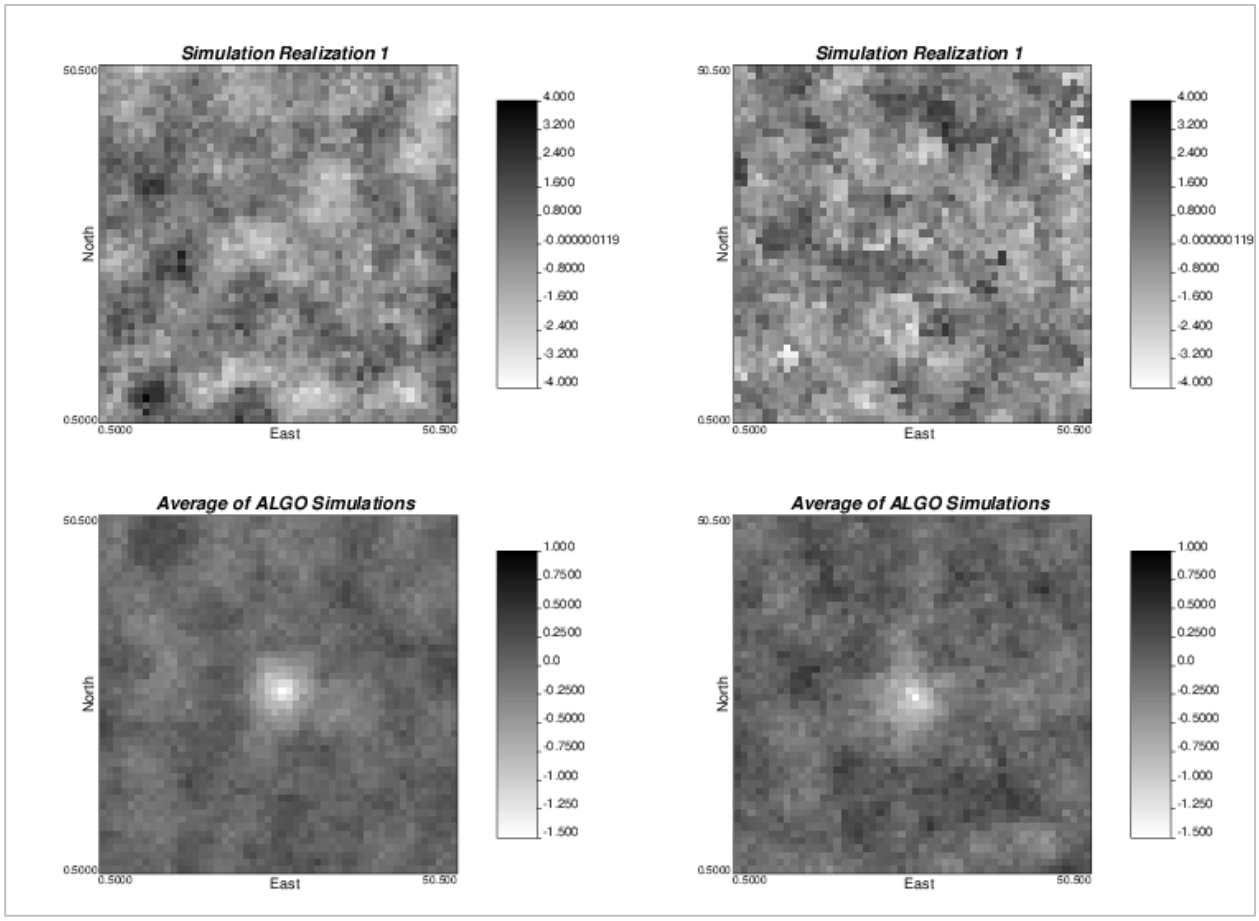


Figure 7: Comparison of SGS and SIA Conditional Variance

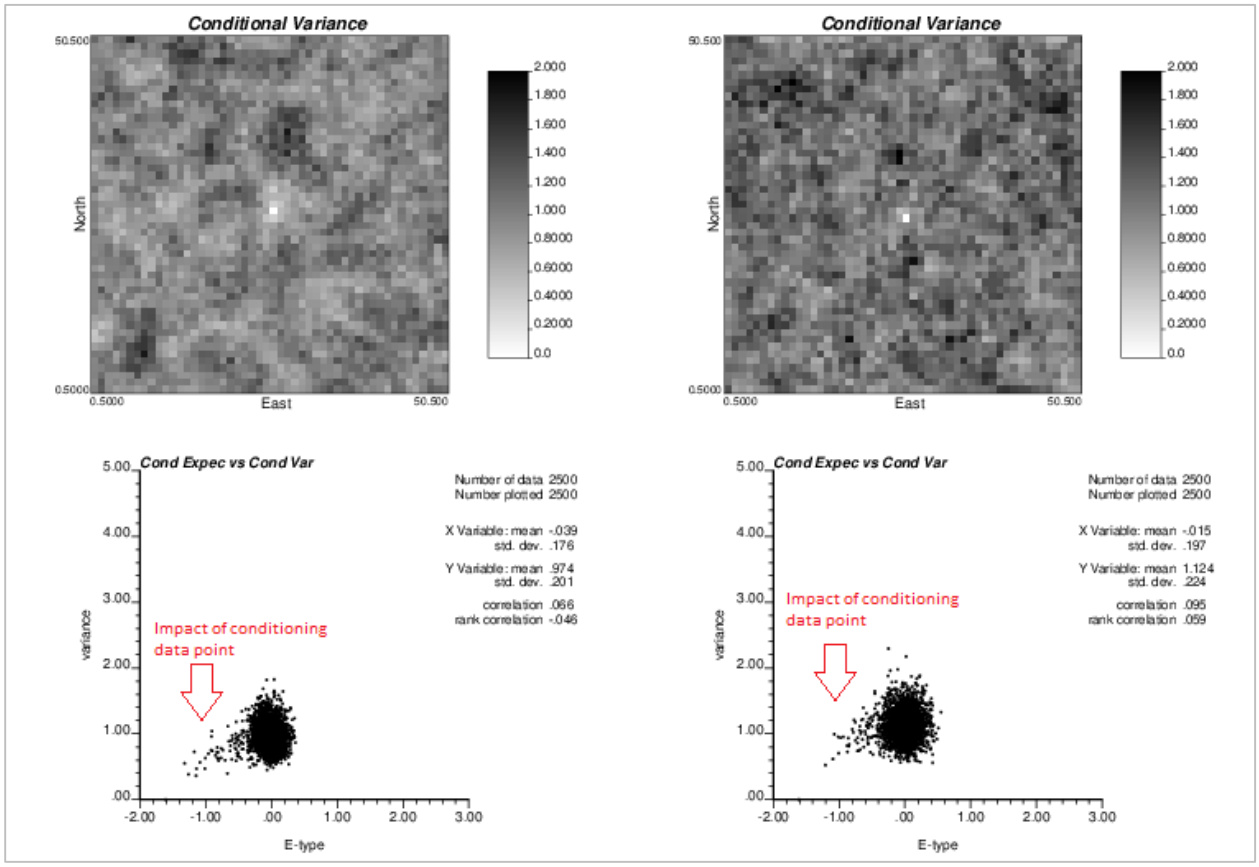


Figure 8: Comparison of SGS and SIA Experimental Variograms

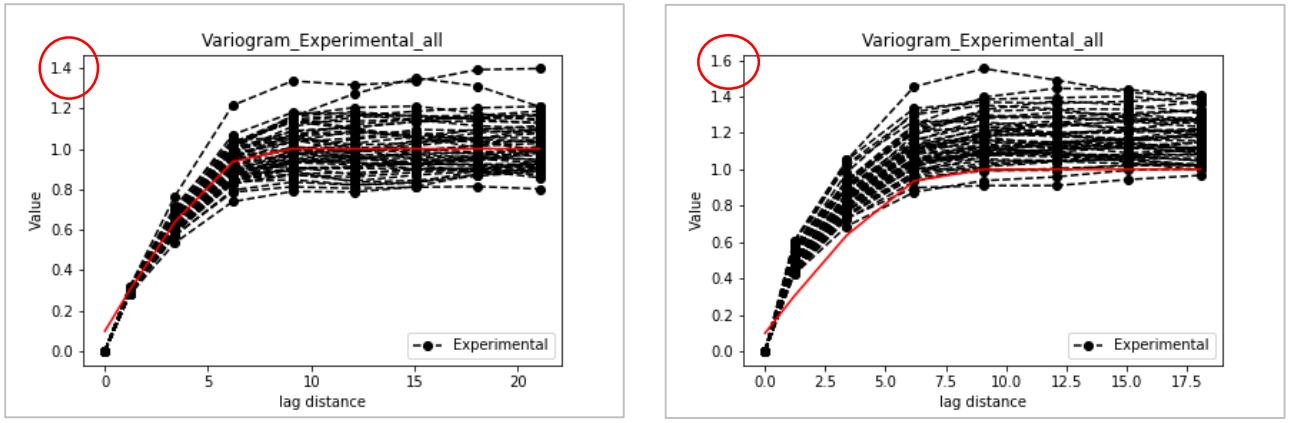
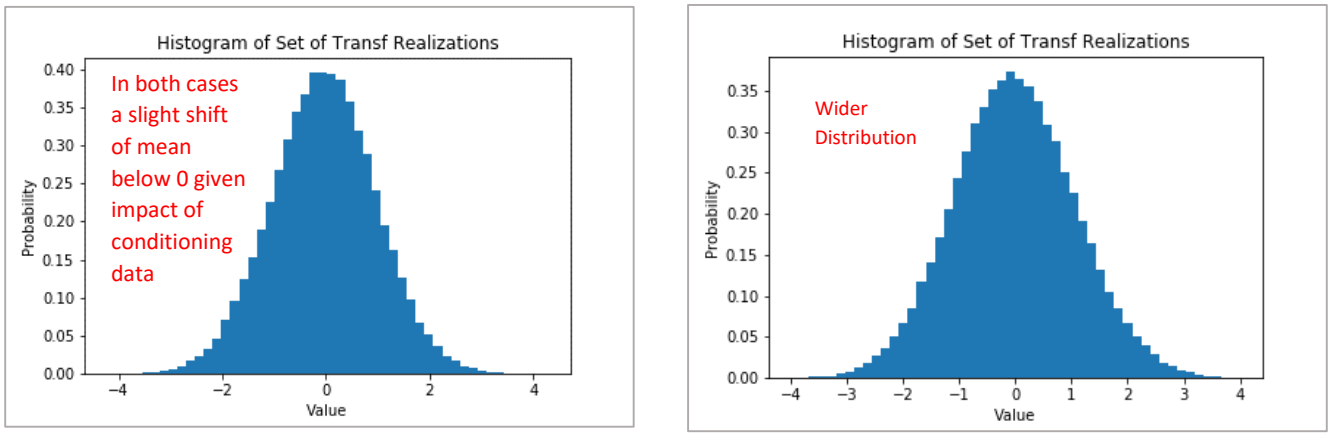


Figure 9: Comparison of SGS and SIA Histograms from Simulation Test



The simulation results back-transformed with lognormal reference distribution to be raw values are shown below with SGS results on the left and SIA results on the right in figures 10, 11 and 12.

Figure 10: Comparison of SGS and SIA Raw Realizations and Average (E-Type)

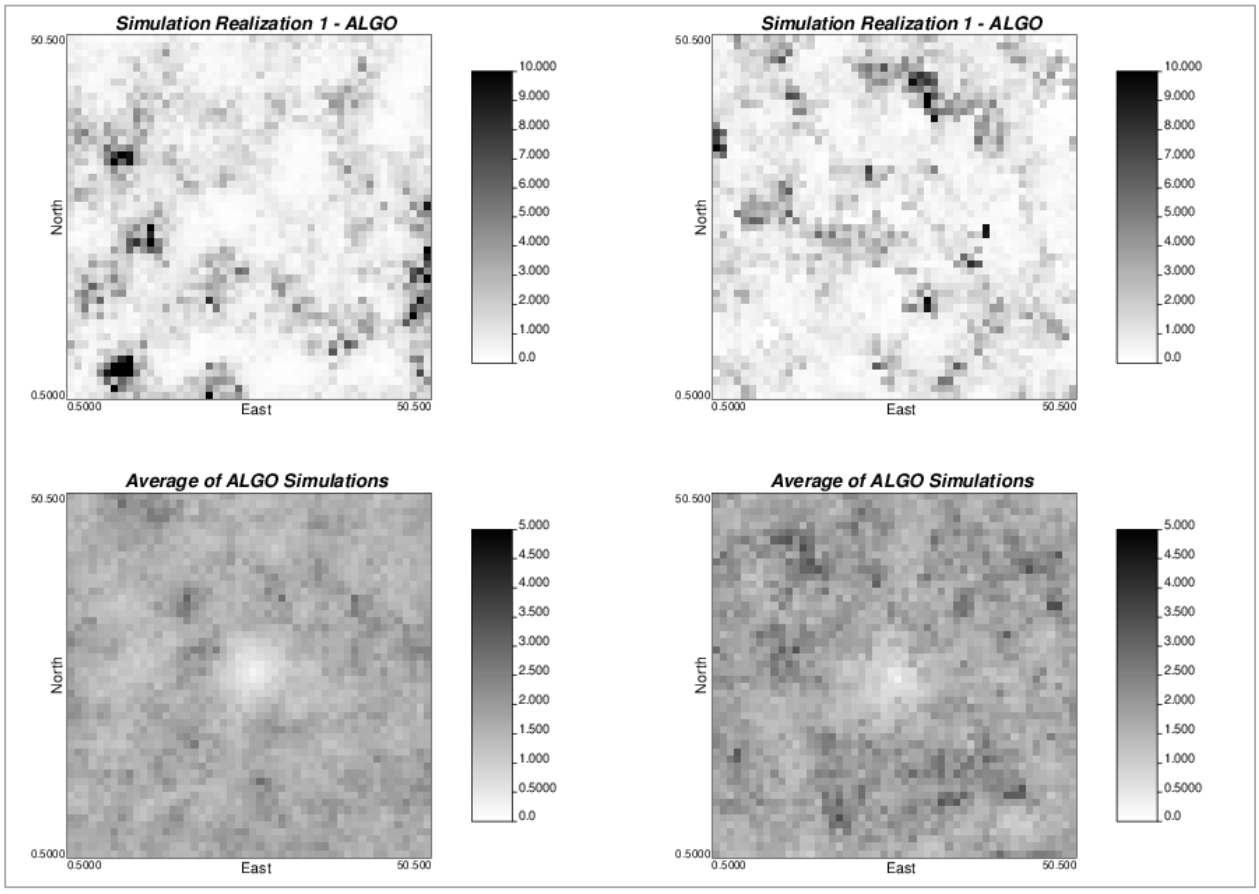


Figure 11: Comparison of SGS and SIA Raw Conditional Variance

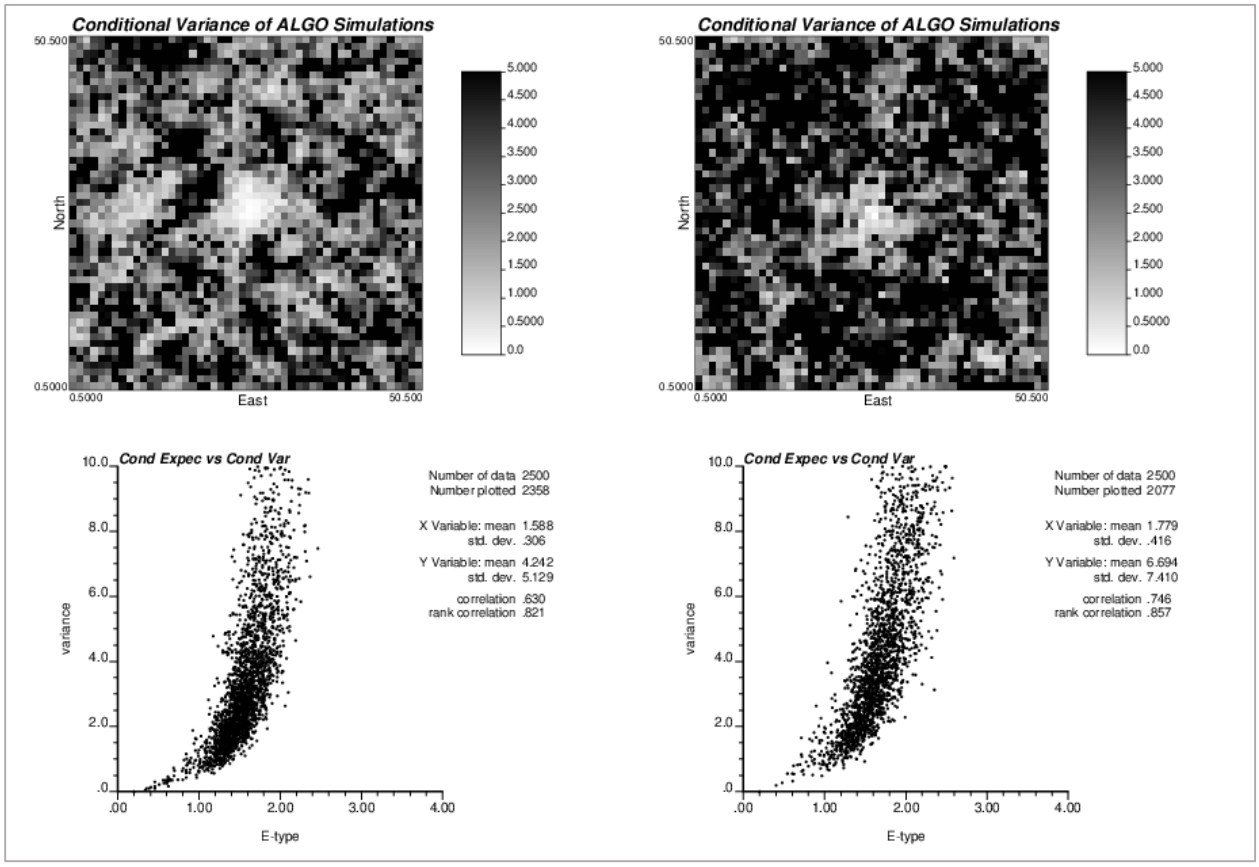
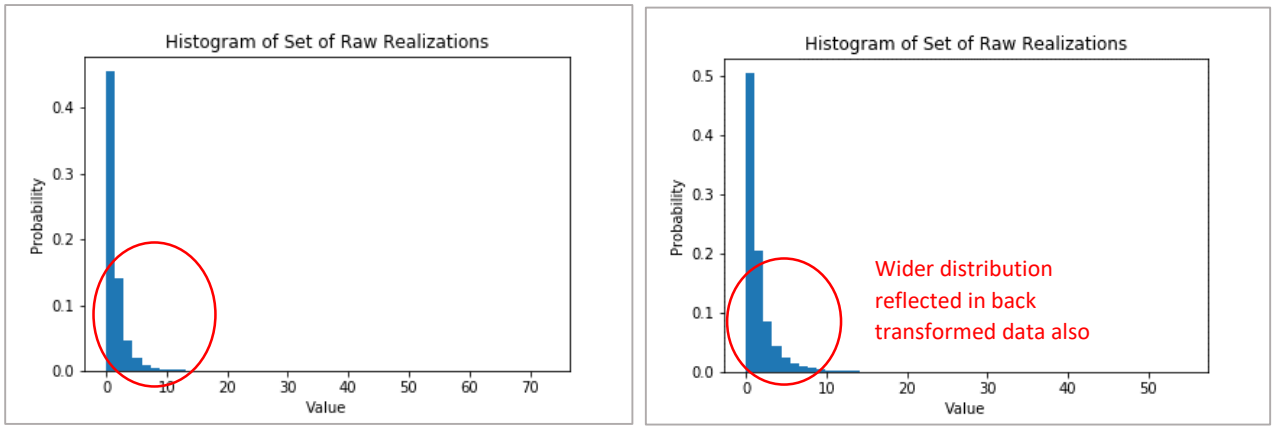


Figure 12: Comparison of SGS and SIA Raw Histograms



A second experiment was run using a larger variogram range. The results showed increased continuity in both SIA and SGS results as shown below.

Figure 13: Comparison of SGS and SIA Simulations with Increased Variogram Range; Realizations and Average (E-Type)

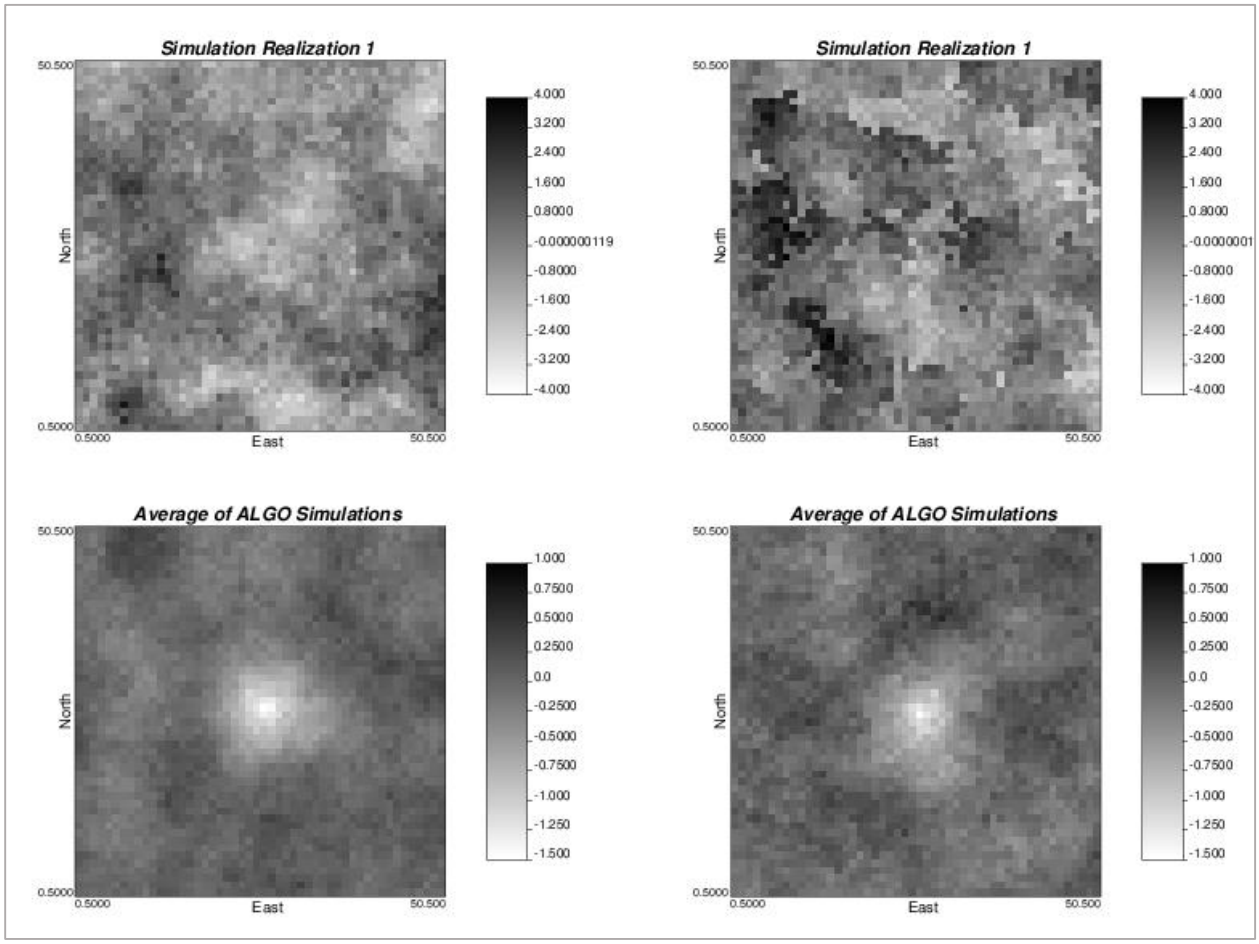
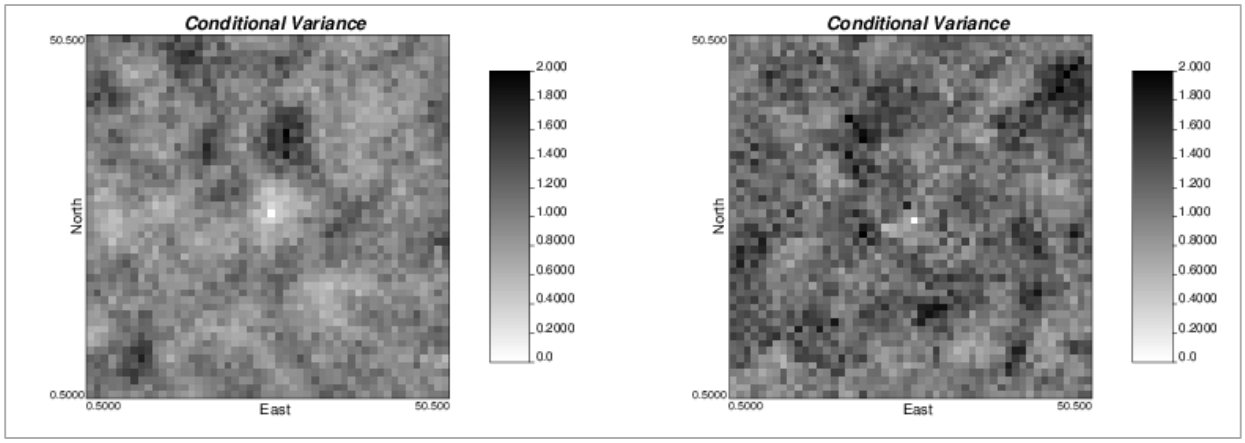


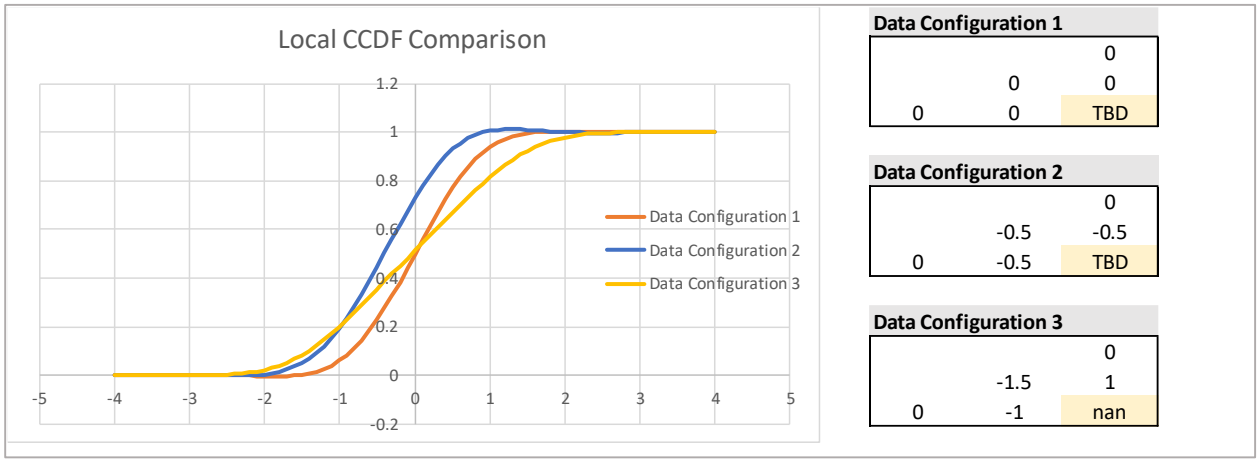
Figure 14: Comparison of SGS and SIA Simulations with Increased Variogram Range Conditional Variance



The cause of variability increases in the SIA realizations relative to the SGS realizations is described further. The local CCDF form under SIA is impacted by both the mean and the variance of the conditioning data. This is illustrated below using the SIA algorithm. First a local CCDF is created conditioned to a uniform set

of zero values, second a uniform value of negative 0.5 is used for all adjacent conditioning points (it can be seen below in figure 15 that SIA behaves like SGS and simply shifts the curve to the new mean). Finally, a curve with the same local mean as the first curve (0.0) but a much higher variance is shown to result from conditioning data with the same zero mean but much higher variance.

Figure 15: Comparison of SGS and SIA Simulations with Increased Variogram Range Conditional Variance



Data Configuration 1

		0
	0	0
0	0	TBD

Data Configuration 2

		0
	-0.5	-0.5
0	-0.5	TBD

Data Configuration 3

		0
	-1.5	1
0	-1	nan

3.3. Examining Variance in SIA Results

A further examination of the conditional variance relative to various model input parameters was conducted.

Base Results (50 x 50 grid)

- Variogram Range and search radius of 12
- Conditioning data point -1.6 normal score (0.2 lognormal)
- 50 simulations

Figure 16: Conditional variance of 50 x 50 simulation with single conditioning point showing gradual increase of conditional variance away from conditioning data point

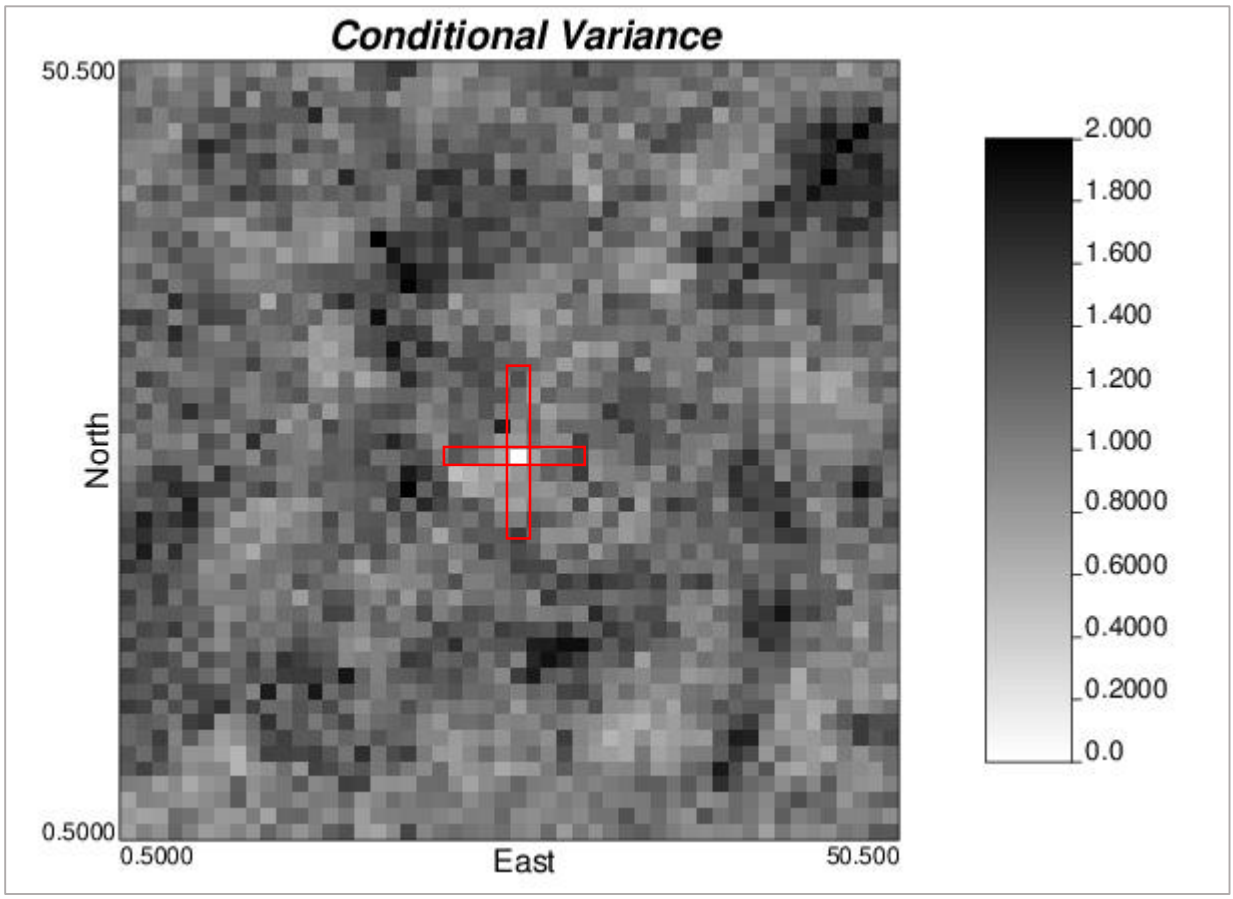
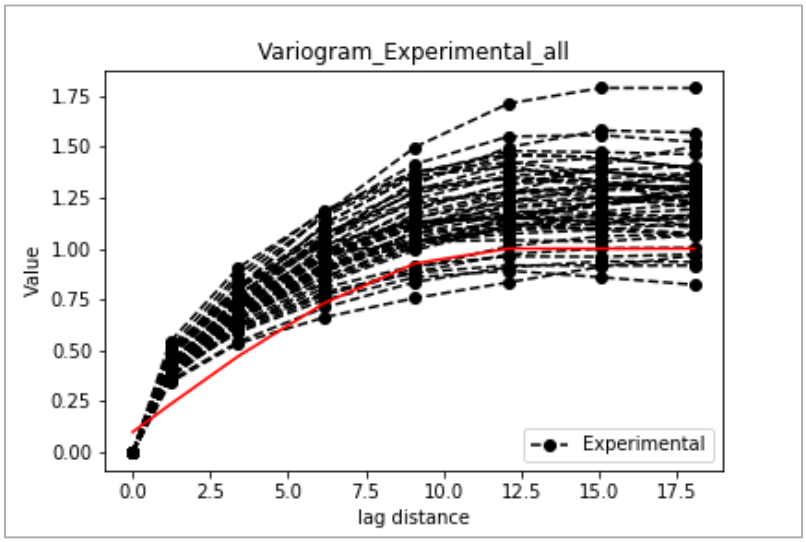


Figure 17: Experimental Variograms of 50 x 50 simulation with single conditioning point



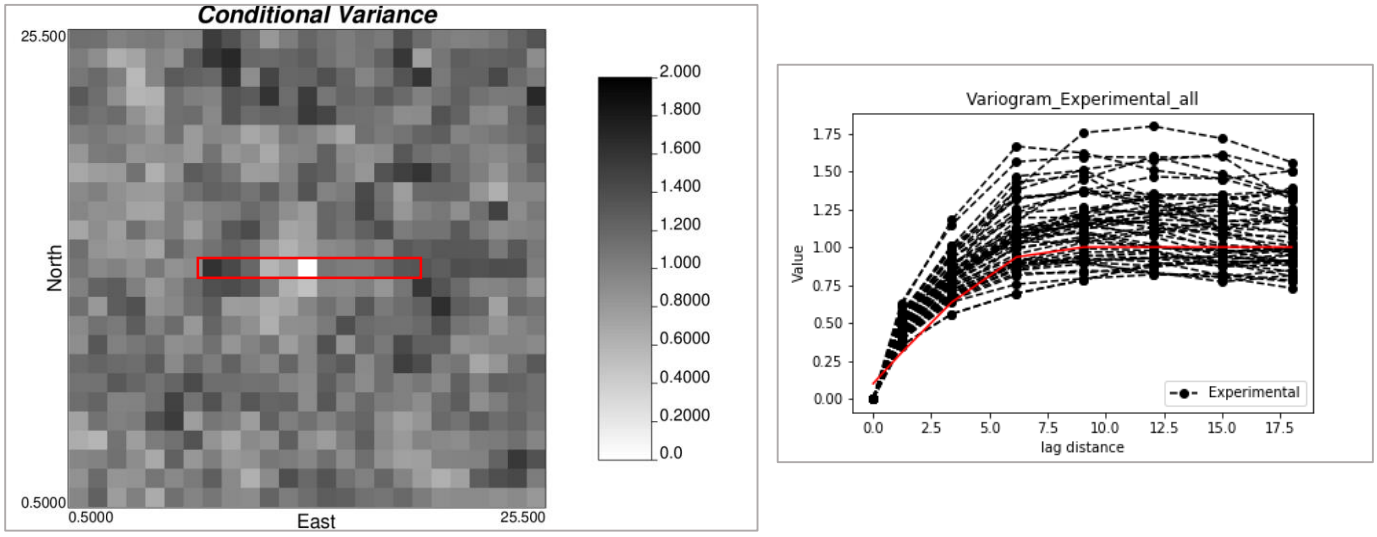
A number of other scenarios were examined on 25 x 25 grids to review impact of assumptions on variance as shown in map of conditional variance and experimental variograms for the set of realizations.

Hypothesis 1: less smooth conditional variance around the conditioning data point in SIA driven by the fact that starting with a low value creates greater dispersion of samples and a broader curve. This also drives higher overall variance in the realizations

Set Up 1

- Variogram Range and search radius of 8
- Conditioning data point -1.6 normal score (0.2 lognormal)
- 50 simulations

Figure 18: Conditional variance map and experimental variograms for set up 1

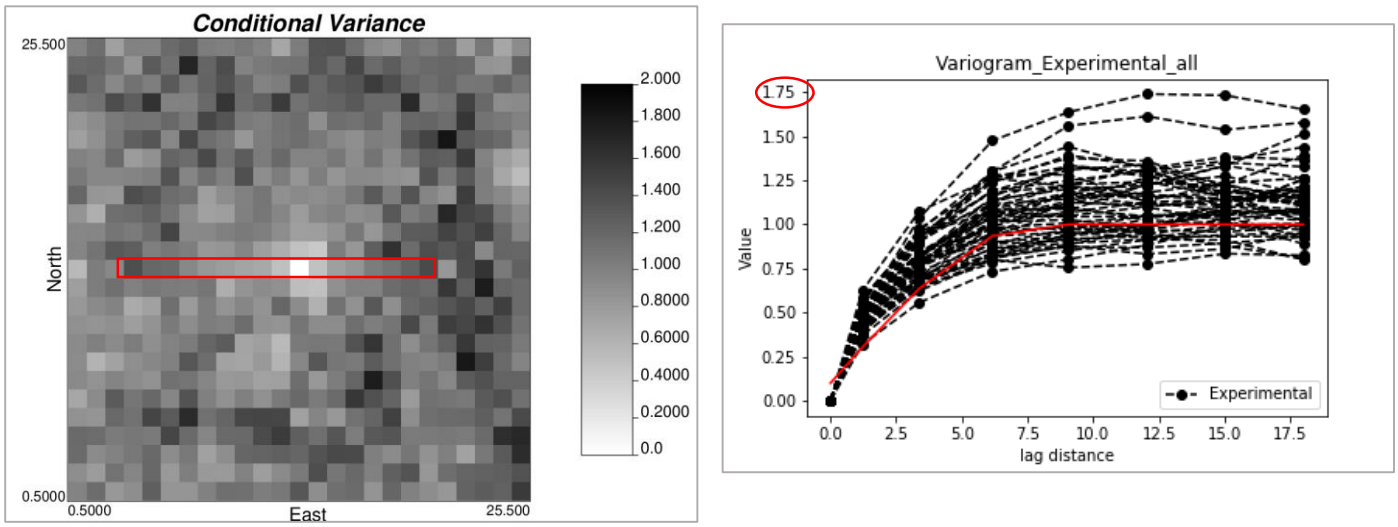


Set Up 2

- Variogram Range and search radius of 8
- Conditioning data point 0.0 normal score (1.0 lognormal)
- 50 simulations

Relative to set up 1, the use of a central conditioning data point instead of a low value conditioning data point smooths the conditional variance around the conditioning data point as anticipated. Overall variability also appears reduced slightly in the variograms.

Figure 19: Conditional variance map and experimental variograms for set up 2 showing reduced variance from non-extreme conditioning values

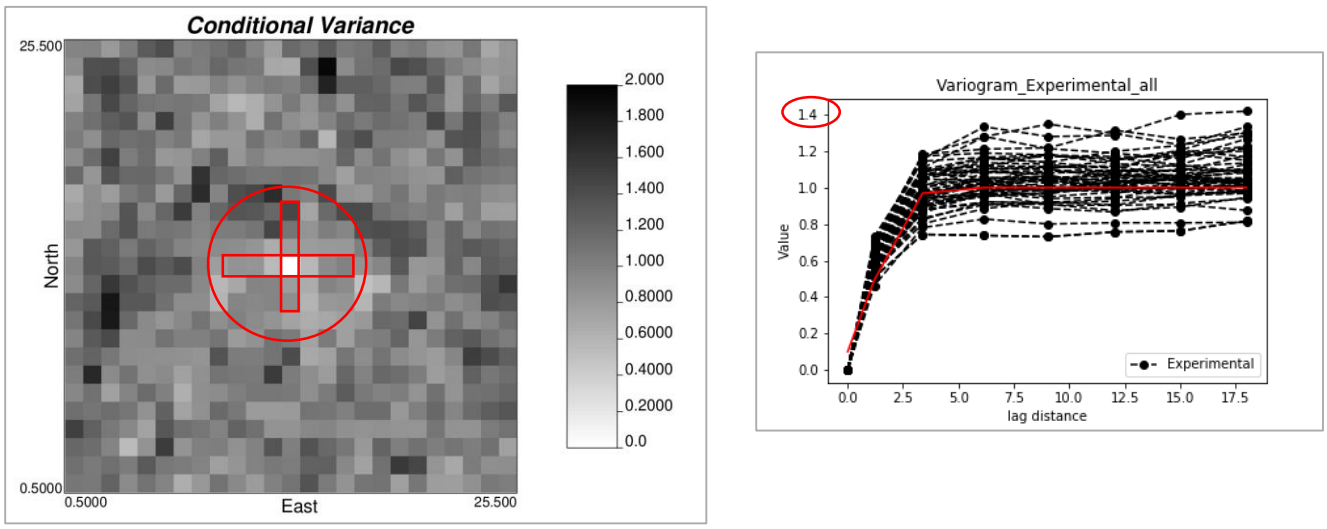


Hypothesis 2: Variogram range should be correlated with variability. A short variogram range should result in lower variance (i.e. lower variogram sill) as by definition, conditioning data variance at any given point likely to be lower (fewer widely spaced values expected in the covariance matrix).

<u>Set Up 3</u>

- <u>Variogram Range and search radius of 4</u>
- Conditioning data point 0.0 normal score (1.0 lognormal)
- 50 simulations

Figure 20: Conditional variance map and experimental variograms for set up 3 showing reduced variance from lower search radius and variogram range



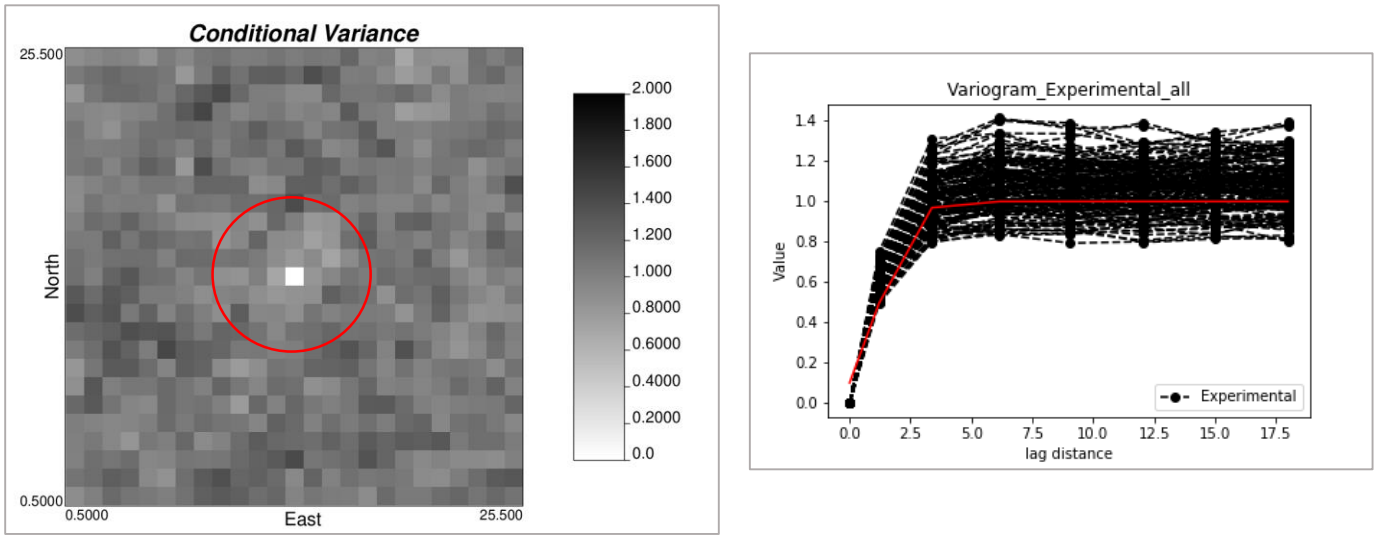
Relative to set up 2, the use of a shorter variogram range reduces the variance in each realization as shown in the experimental variograms that cluster more evenly around the modelled variogram. Conditional variance is also smooth over a shorter range given the lower variogram range.

Hypothesis 3: increasing number of simulations should smooth conditional variance map

Set Up 4

- Variogram Range and search radius of 4
- Conditioning data point 0.0 normal score (1.0 lognormal)
- 150 simulations

Figure 21: Conditional variance map and experimental variograms for set up 4

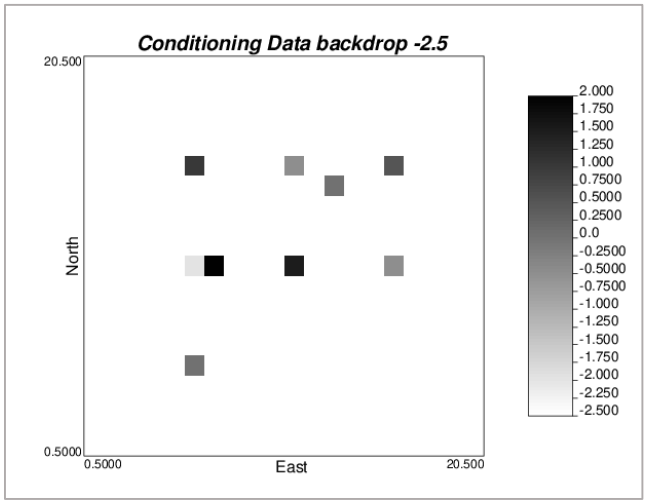


Relative to set up 3, the use of a greater number of simulations significantly smooths the conditional variance map.

3.4. Comparison of Random Function Models

A blank 20 x 20 array was initialized. Nine conditioning data points were assigned. The conditioning data is shown below.

Figure 22: Conditioning data for random function model comparison



The SIA simulation algorithm was then run over the conditioning data for four scenarios:

A) Pure Diffusive Gaussian Model
B) Pure Diffusive Gamma Model (Shape of 2)
C) Gaussian Model with Destructuration Coefficient of 0.0001
D) Gamma Model with Destructuration Coefficient of 0.0001 (Shape of 2)

In each case, 50 simulations were completed producing 50 realizations of the gamma or Gaussian score data and 50 realizations of the back transformed raw data. A simple spherical variogram model was assumed with range of 5 units and 0.1 nugget (total sill is 1.0). The results for the four cases are shown below in figure 23 to 29.

Figure 23: Simulation realizations from four random function model assumptions; top left is Gaussian diffusive, top right is gamma diffusive, bottom left is Gaussian non diffusive and bottom right is gamma non diffusive.

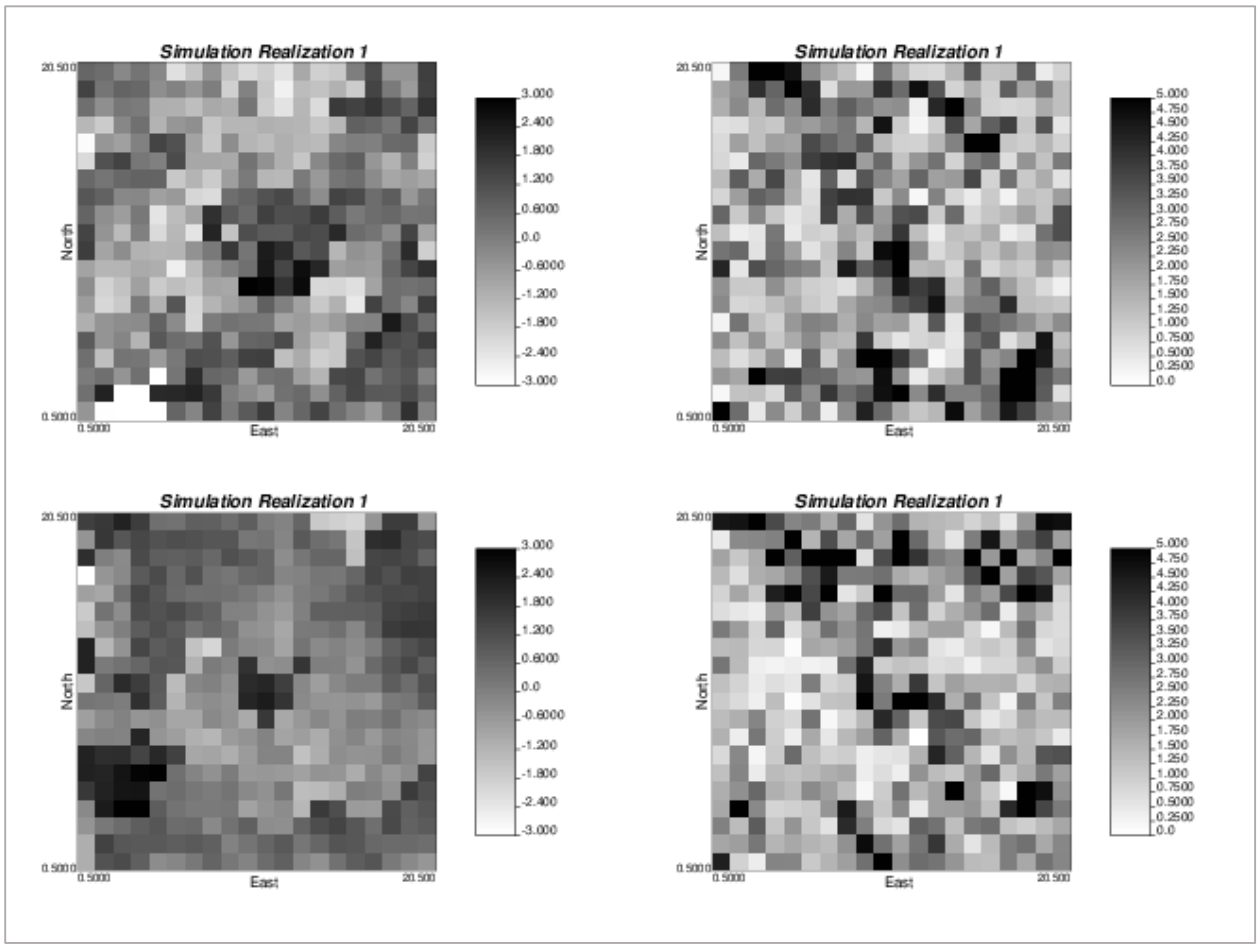


Figure 24: Simulation realization average (E-type) from four random function model assumptions; top left is Gaussian diffusive, top right is gamma diffusive, bottom left is Gaussian non diffusive and bottom right is gamma non diffusive.

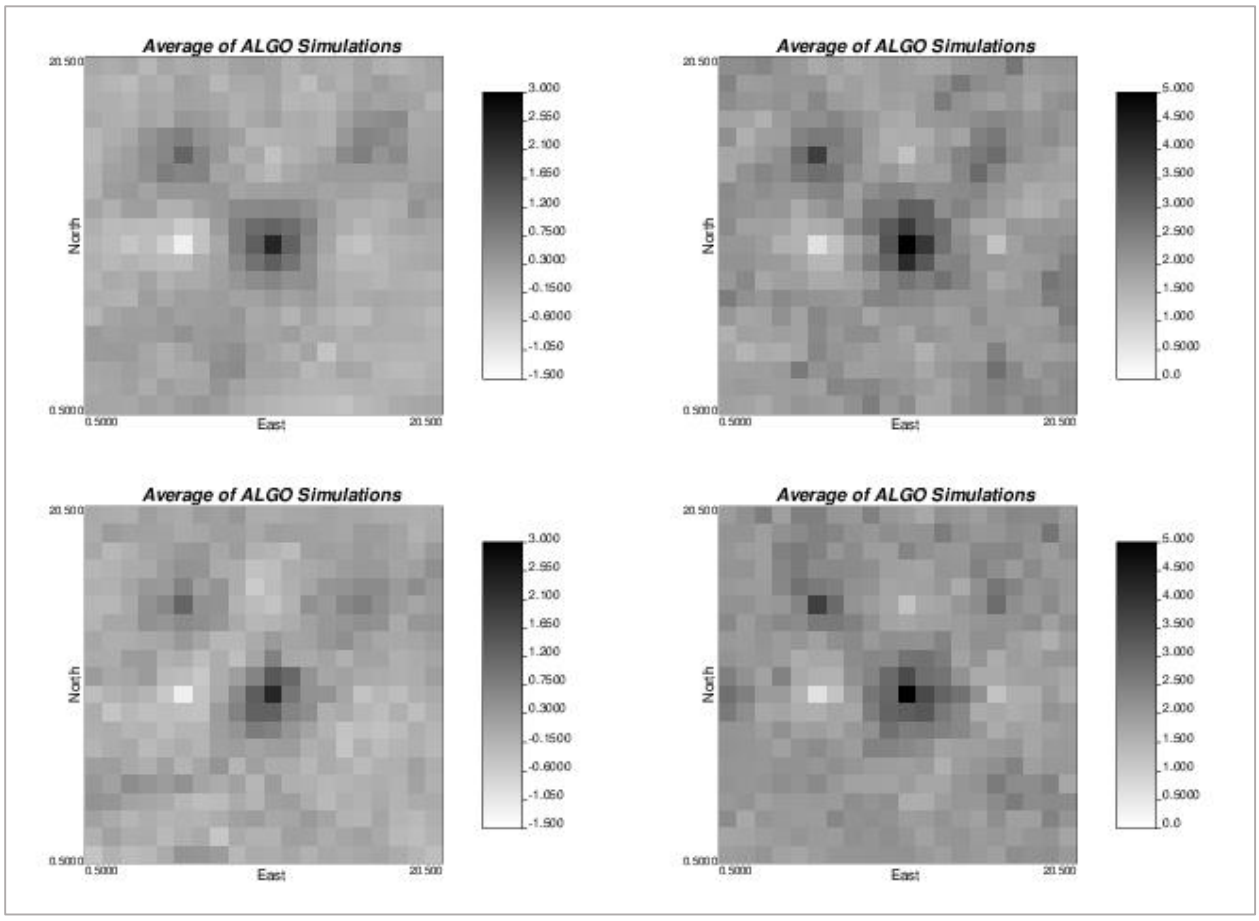


Figure 25: Simulation realization histograms from four random function model assumptions; top left is Gaussian diffusive, top right is gamma diffusive, bottom left is Gaussian non diffusive and bottom right is gamma non diffusive.

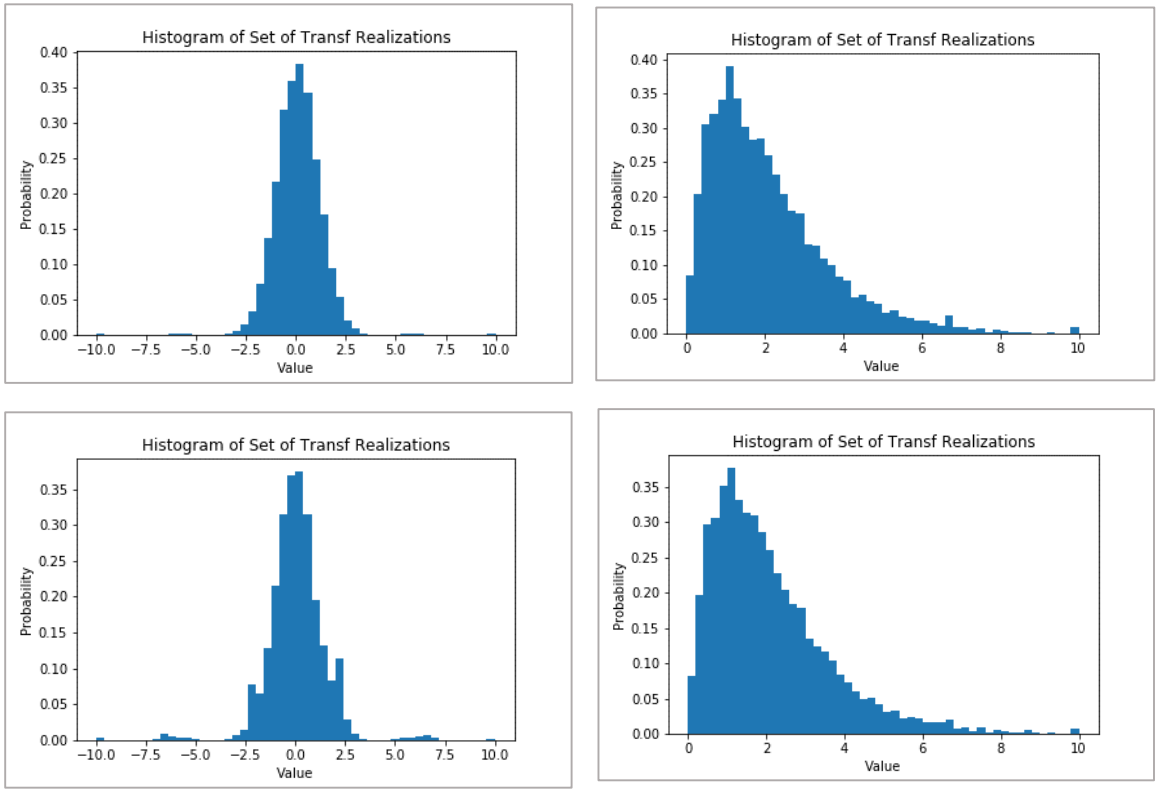


Figure 26: Conditional variance from four random function model assumptions; top left is Gaussian diffusive, top right is gamma diffusive, bottom left is Gaussian non diffusive and bottom right is gamma non diffusive.

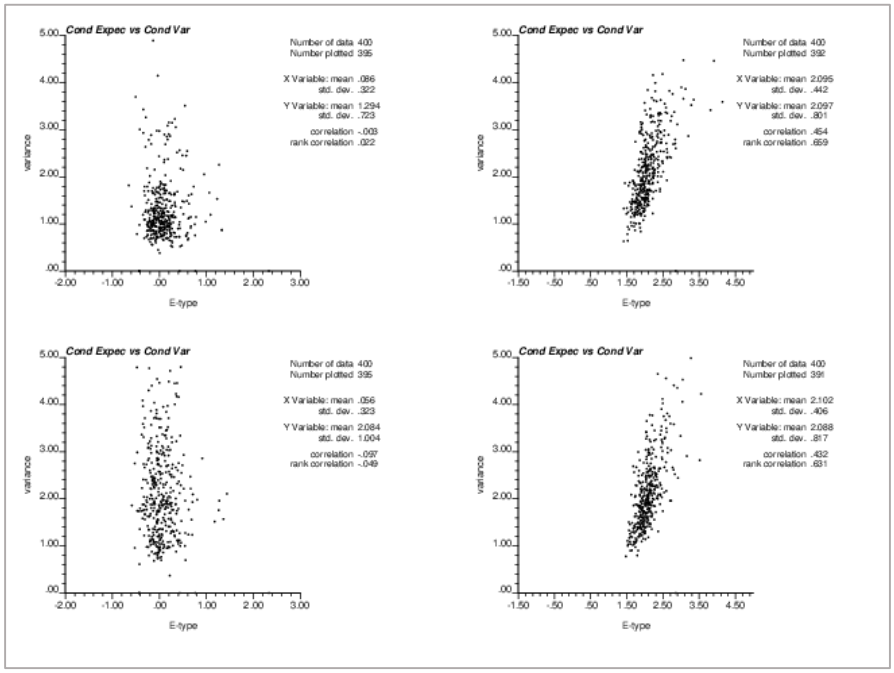


Figure 27: Back transformed average (E-type) of 50 simulations from four random function model assumptions; top left is Gaussian diffusive, top right is gamma diffusive, bottom left is Gaussian non diffusive and bottom right is gamma non diffusive.

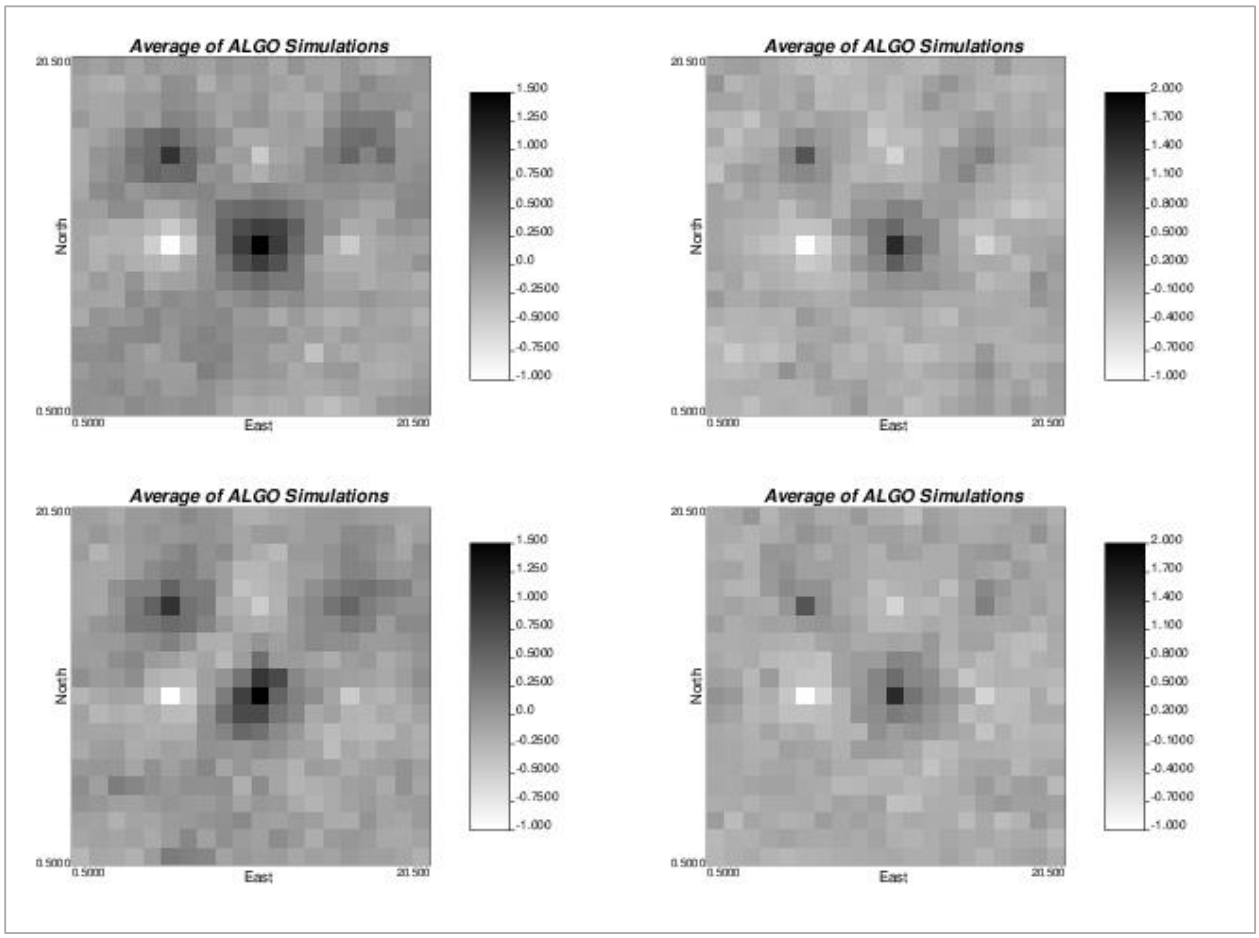


Figure 28: Back transformed conditional variance from four random function model assumptions; top left is Gaussian diffusive, top right is gamma diffusive, bottom left is Gaussian non diffusive and bottom right is gamma non diffusive.

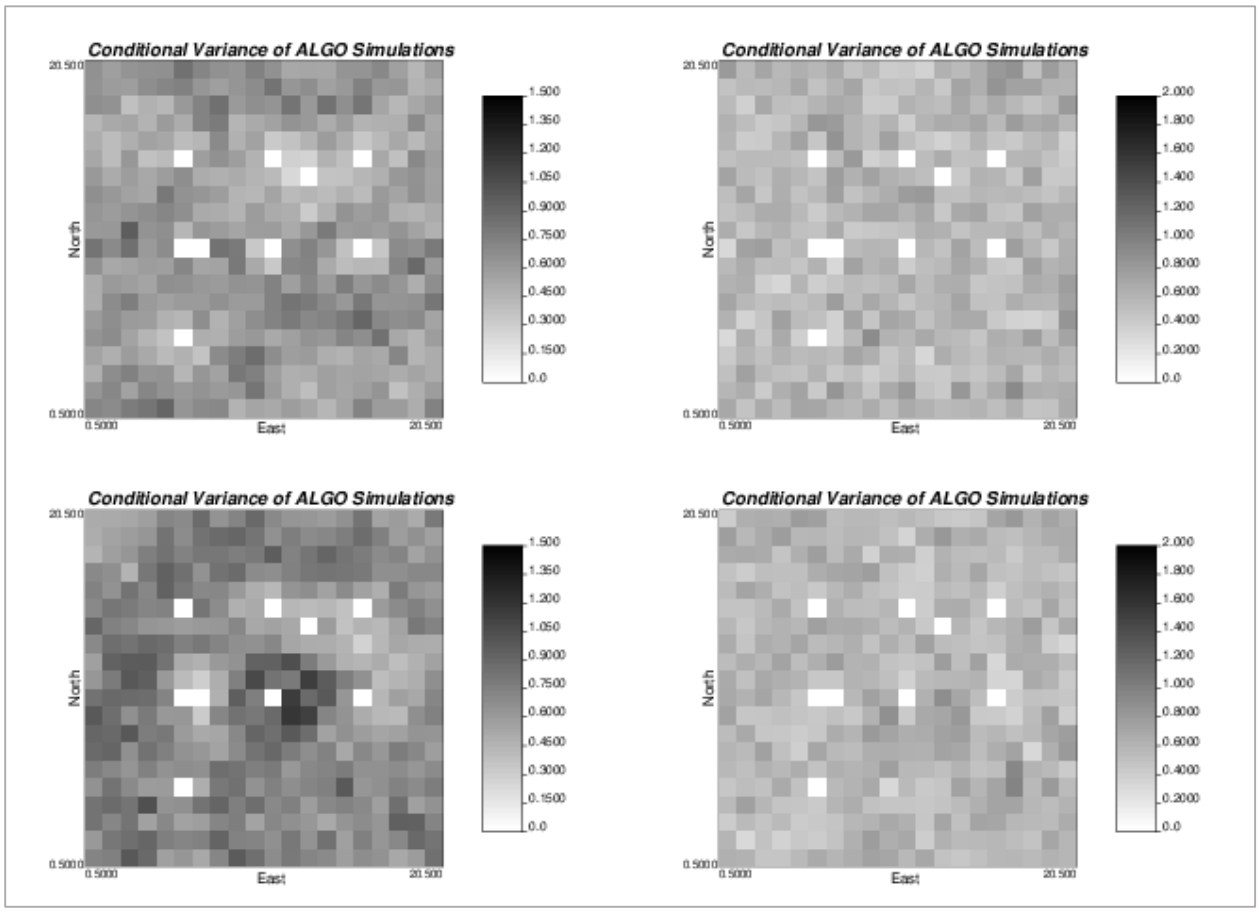
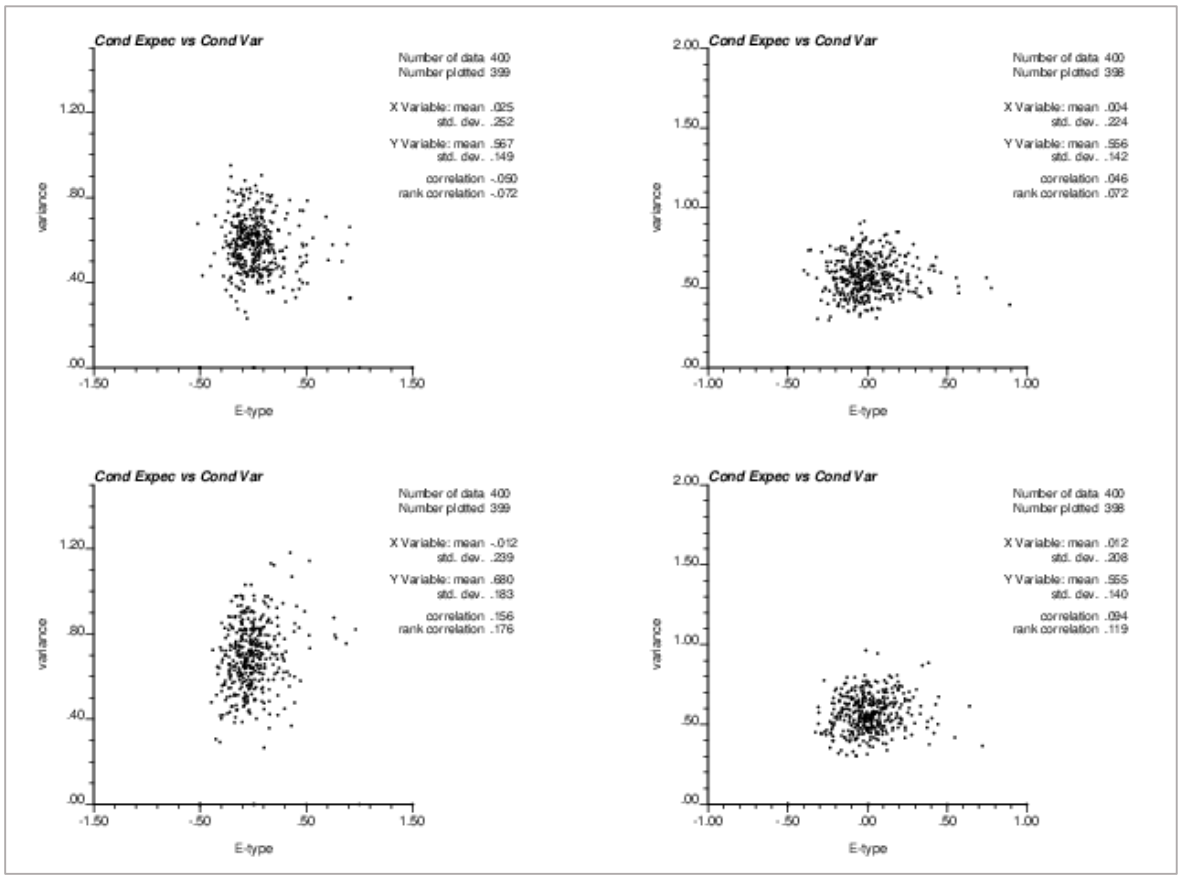


Figure 29: Back transformed conditional variance vs. average (E-Type) from four random function model assumptions; top left is Gaussian diffusive, top right is gamma diffusive, bottom left is Gaussian non diffusive and bottom right is gamma non diffusive.

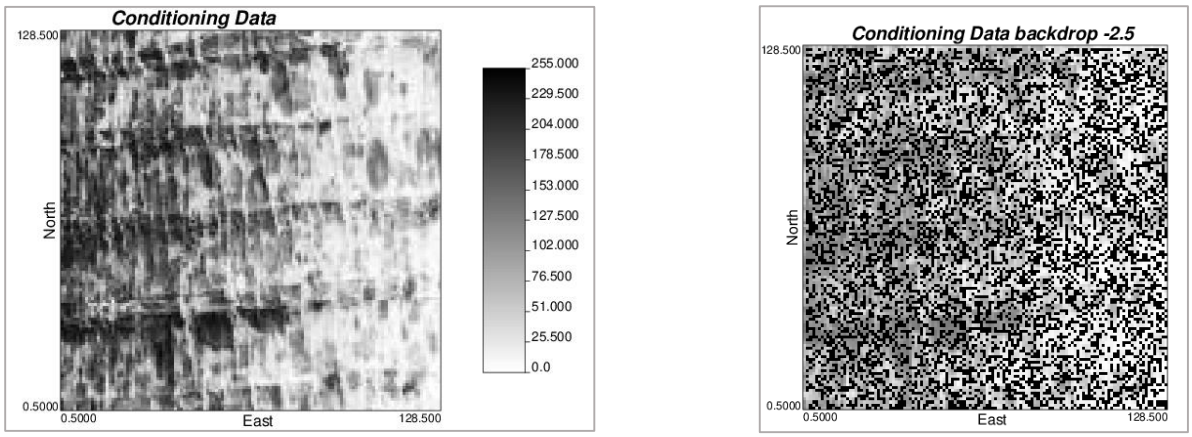


The Gamma model with shape two has a CDF with a lower slope than pure Gaussian, so in general has slightly more variability in simulation. Additionally, conditioning data (in particular extreme conditioning data) bends the local CCDF curve outward creating a more random (variable) simulation draw, vs the Gaussian curve which not only bends outward to accommodate extreme conditioning data but also recenters itself based on the conditioning data. The result is higher conditional variance at higher values for Gamma models as shown in scatterplot of E-Type vs conditional variance. For similar reasons described above, in situations with extreme conditioning data, the local gamma CCDF curve becomes quite variable (i.e., low slope) which reduces the impact of tightening the corners of the local CCDF to account for clustering (relative to a Gaussian local CCDF that translates horizontally to account for extreme conditioning data).

3.5. Illustrative Analysis of a Sampled Data Set

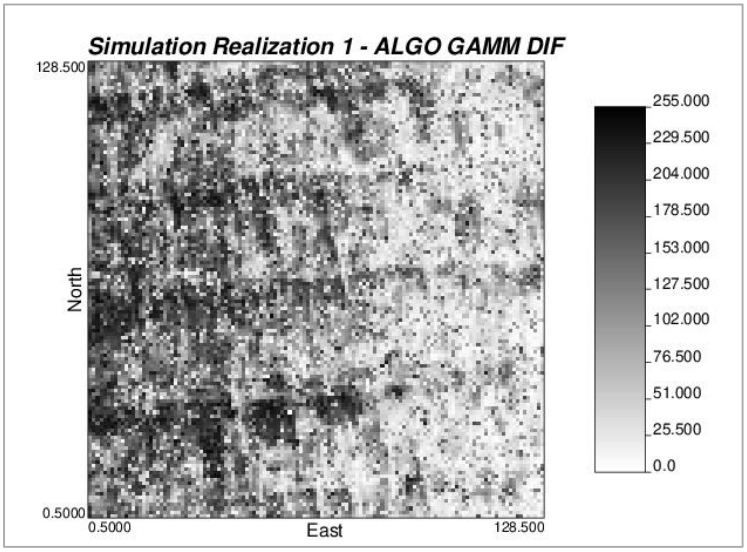
A raw 128 x 128 grey scale image was selected and randomly sampled as shown below in figure 31.

Figure 30: Original Grey Scale Data Set and Randomly Sampled Data Set



A single simulation realization was run using an assumed spherical model with range of 5 and nugget of 0.1. The simulation was run using a gamma model and a destructuration coefficient of 0.00001. The search radius was set to 4 and the minimum points was set to 1.

Figure 31: Single Simulation Realization of Unsampled Values Using SIA Algorithm



4. Conclusions

The SIA algorithm produces intuitive results that deliver expected behavior in both histogram and variogram as well as conditional variance. The algorithm produces distinct results for Gaussian and gamma models and clearly shows the impact of destructuration through clustering of values and sharp boundaries. These initial results suggest that SIA may provide an effective tool to capture a greater range of random function models in simulation and measurement of uncertainty.

5. Acknowledgments

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2017-04200 and RGPAS-2017-507956.

6. References

Bhattacharya R., (2010) Normal Approximation and Asymptotic Expansions Society for Industrial and Applied Mathematics

Chiles J.P., Delfiner P., (1999) Geostatistics Modelling Spatial Uncertainty Wiley, New York NY

Dimitrakopoulos, M. (2009). High-order Statistics of Spatial Random Fields: Exploring Spatial Cumulants for Modeling Complex Non-Gaussian and Non-linear Phenomena. Mathematical Geosciences, 42(1), 65–99

Edgeworth F., (1907) On the Representation of Statistical Frequency by a Series. Journal of the Royal Statistical Society, 70 (1), 102–106

Emery X., (2002) Conditional simulation of nonGaussian random functions, Mathematical Geology, 34, p 79-100

Emery X., (2005a) Variograms of Order ω: a tool to validate a bivariate distribution model, Mathematical Geology, 37, p. 163-181

Emery X., (2005b) Geostatistical Simulation of Random Fields with Bivariate Isofactorial Distributions by Adding Mosaic Models. Stochastic environmental research and risk assessment, 19 (5), 348–360

Emery X., (2005c) Conditional Simulation of Random Fields with Bivariate Gamma Isofactorial Distributions. Mathematical geology, 37 (4), 419–445

Emery X., (2006) A Disjunctive Kriging Program for Assessment Point-Support Conditional Distributions Computers and Geosciences 32 965-983

Emery X., (2007) Using the Gibbs Sampler for Conditional Simulation of Gaussian-Based Random Fields. Computers & geosciences, 33 (4), 522–537

Emery X., (2008) Substitution Random Fields with Gaussian and Gamma Distributions: Theory and Application to a Pollution Data Set. Mathematical geosciences, 40 (1), 83–99

Emery X., Kremer F., (2008) A Survey of Random Field Models for Simulating Mineral Grades, in VII International Geostatistics Congress.

Emery X., Ortiz J.M. (2004) Shortcomings of multiple indicator kriging for assessing local distributions, Applied Earth Science (Trans. Inst. Min. Metall. B), 113(4): 249-259.

Goovaerts, P. (1997) Geostatistics for Natural Resources Evaluation, Oxford University Press, New York

Machuca-Mory D. F., Ortiz J. M., and Deutsch, C. V., (2008) On the challenge of using sequential indicator simulation for the estimation of recoverable reserves, International Journal of Mining, Reclamation and Environment, 22(4): 285-299

Mustapha D., Dimitrakopoulos R., (2010) Generalized Laguerre Expansions of Multivariate Probability Densities with Moments. Computers & mathematics with applications 60 (7), 2178–2189

Mustapha D., Dimitrakopoulos R., (2010b) High-order Stochastic Simulation of Complex Spatially Distributed Natural Phenomena. Mathematical geosciences. [Online] 42 (5), 457–485.

Ortiz J., (2004) A Step by Step Guide to Bi-Gaussian Disjunctive Kriging. In Geostatistics Banff; Springer Netherlands: Dordrecht; Vol. 14, pp 1097–1102

Ortiz J., (2019) MINE 834 Lectures. Queens University

Skovgaard M., (1986) On Multivariate Edgeworth Expansions International Statistical Review / Revue Internationale De Statistique, vol. 54, no. 2, pp. 169–186

Taylor Polynomial Expansions Website https://mathinsight.org/applet/taylor_polynomial accessed January 2021

Wallace D.L., (1958) Asymptotic Approximations to Distributions, The Annals of Mathematical Statistics, vol. 29, no. 3, pp. 635–654

Wilson G. A., Wragg A., (1973) Numerical methods for approximating continuous probability density functions using moments IMA J. Appl Math. 12 165173

Withers N., (2014) The Dual Multivariate Charlier and Edgeworth Expansions Statistics & probability letters 87, 76–85

Yao D., Dimitrakopoulos R., (2018) A New Computational Model of High-Order Stochastic Simulation Based on Spatial Legendre Moments, Mathematical geosciences 50 (8), 929–960

Defining geological units using geochemical data and unsupervised machine learning[1]

Noble E. Potakey (n.potakey@queensu.ca)
Julian M. Ortiz (julian.ortiz@queensu.ca)

Abstract

One of the preliminary and arguably the most crucial step in a mineral resource evaluation campaign is the determination of the geological domains. Conventional geological methods establish domains primarily by grade zoning or spatial clustering techniques. Even though information about the geology is recorded, most model-based domains do not make much of the geological information derived from logging. Domains are best established with the geological information supported by an accurate statistical analysis of the geochemical data and a good understanding of the deposit. The advent of machine learning techniques such as cluster analysis has advanced this course by providing algorithms that can handle large volumes of multivariate data and try to reproduce geological domains. This paper shows the application of a model-based cluster analysis as a machine learning tool to an exploratory drill hole data set from an undisclosed copper porphyry deposit. The K means algorithm, which was applied in this study utilizes the continuous nature of the non-categorical variables to establish domains. The algorithm generated spatial clusters which had some correlation with the alteration unit even though a confusion matrix revealed the flaws of the method in misclassifying most of the geological units. The choice of the most appropriate number of clusters (domains) to be formed, as well as the selection of variables to drive the clustering process can be challenging when performing k means clustering, and the appraisal of an expert is still necessary, as the results are subjective.

1. Introduction

The mining world has not had enough of mineral exploration. Geological mapping, geophysical investigation, sampling of outcrops, logging of drill cores are examples of exploratory data that needs to be analyzed leaving geologists and engineers overwhelmed with large number of variables. Among these data collected are categorical variables about the lithology, alteration and mineralization of an ore body which is largely obtained from core logging and after measuring several physical, chemical, and mineralogical properties of the rock (Bosch et al., 2002). Knowledge about these geological units is important because domains are traditionally established based on them. Domaining in mineral resource evaluation is a big step in mining because it serves as a backbone for subsequent geostatistical estimation and simulation of the ore body. A poor classification of these domains can lead to the mixing of populations, which can result in bad resource estimates, endangering the valuation of grades and tonnages (Emery & Ortiz, 2005). A simulated model of an ore body will be a success or failure depending

[1] Cite as: Noble E. Potakey, Julian M. Ortiz (2022) Defining geological units using geochemical data and unsupervised machine learning, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-03, 47-58.

on the accuracy of the domains established. A domain is formed when an ore body is partitioned into groups of similar characteristics. For example, groups of high element concentration or groups of low rock hardness can form a domain.

Domains are best established with the geological information derived from logging supported by an accurate statistical analysis of the geochemical data and a good understanding of the deposit. However, in most model-based domains, attention is not really given to the geological information when characterizing different mineralized zones even though it is proven that ore grades vary in relation to changes in the geological properties such as mineralogy, lithology and alterations (Yasrebi et al., 2013). Despite the non-reliance of geological information to create domains in most mining environments, it remains one of the fundamental information to building a good domain (Sterk et al., 2019). In the context of unsupervised machine learning, cluster analysis emerges as an efficient tool in classifying sample points based on the intrinsic properties of the input variables. K means algorithm clusters data by grouping sample into clusters of equal variances thereby minimizing a phenomenon known as the *inertia* or within-cluster sum-of-squares (Adams, 2018). This results in clusters that contain objects with similar features and at the same time different from objects belonging to a different cluster. Although the algorithm clusters based on statistical parameters, knowledge about the deposit was used to select variables to drive the clustering process based on their significance to the geological units. For instance, most of the variables in our data were attached to the alteration type and hence variables that are trace elements to the porphyry copper deposit were selected for clustering (Mg, Al, Ga, Li, SC, V). This was done to derive a fast, better performing, and easy to understand model.

An important factor in K means clustering is the choice of the optimal number of clusters (Moreira et al., 2021). This paper addresses this issue, applying and further discussing some of the methods that can be applied as well as the difficulties found when choosing the best configuration of the clusters. The model is validated by comparing the clustered data with logged geological units. Furthermore, a confusion matrix is computed to analyze the errors of misclassification. It is an expectation that the clustered geochemical data reflects the geology. Proven methods for verifying the spatial relationship of the clusters are rarely mentioned in the literature, other than just applying a visual examination of the results.

The work has been divided in three sections. First, we show the exploratory data analysis for selected variables and their distribution in space. This included the histograms and probability plots of continuous and categorical variables. Secondly, we performed K means clustering of selected variables to drive the clustering based on their relevance to the alteration unit. And finally, the validation of the clustered model.

2. Exploratory Data Analysis (EDA)

Exploration data analysis forms an essential part of this project. The exploratory data was de-surveyed with a 15m run length compositing while breaking intervals by geology for all drill holes using *Vulcan* software. Data analysis was done using the python programming software. The data contained 1898 drill holes from a porphyry copper deposit with 50 continuous variables made of 50 geochemical elements and three categorical variables (lithology, alteration, and mineralogy). A total of 21,416 composite samples with assay results were created with 20, 19 and 25 different lithologies, alterations and mineralogy respectively. Deep samples beyond the depth of 1200m were removed and variables whose concentrations were unaccounted for were also not considered. The following table shows a summary

statistic of some of the variables in the dataset for brevity. The abbreviations “Lito”, “Alt” and “Minz” mean Lithology, Alteration and Mineralization respectively.

Table 1 Summary statistics exploratory data. Note that it does not show all the features of the database.

	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Dhid	21416										
Midx	21416	NaN	NaN	NaN			14971.1				18049.1
Midy	21416	NaN	NaN	NaN			105792.3				107625.6
Midz	21416	NaN	NaN	NaN			1833.9				3110.2
Length	21416	NaN	NaN	NaN			0.04				15
From	21416	NaN	NaN	NaN			0				1197.45
To	21416	NaN	NaN	NaN			0.4				1198.2
Lito	21416	19	50	9603	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Alt	21416	17	51	5621	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Minz	21416	25	70	7712	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Cu_ppm	21416	NaN	NaN	NaN	6112.9	2777.7	46.8	3972	5950	8710	10000
Mo_ppm	21416	NaN	NaN	NaN	92.04652	119.8602	0.45	34.9175	64.7745	113	2890
Mg_ppm	21416	NaN	NaN	NaN	0.442924	0.558493	0.01	0.03	0.094	0.81	3.033
Al_ppm	21416	NaN	NaN	NaN	0.980977	0.661739	0.06	0.46	0.74	1.42	4.519
Ga_ppm	21416	NaN	NaN	NaN	2.591729	2.012614	0.093	1	1.8	3.80525	13.35

Figure 1 below shows the results of the first to third quartiles of all 50 elements to provide us a fair idea of the relationship between the dominant variables and the less dominant ones.

The concentration of copper stands out with extremely high values especially from its third quartile to the maximum. The cumulative probability plot of Cu values shows a fairly log normal distribution with consistently high detection limits of the element as shown in figure 3. The extreme high concentrations of elements could be outliers, measurement or samples errors, or values beyond the detection limit. The negative minimum values of sample 1 (Au) are samples values that were unaccounted for and that was not considered for analysis.

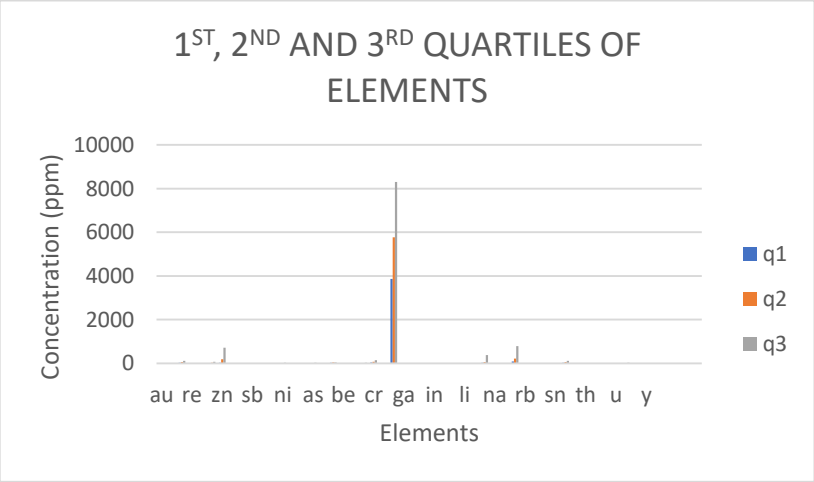


Figure 1 Relatively high Cu concentrations compared other elements.

The following figures represent the box plot of the distribution of copper present in each geological category. This helps us to see the distribution in detail and help identify dominant terrains where our focus should be. Rock code 31, 33 and 50 are the dominant lithologies which host majority of the high grades. Alteration codes 50 and 51 stand out while 50 and 70 are the dominant mineralization codes. Grades of copper are distributed across the features in all three categories and a few outlier values are noted.

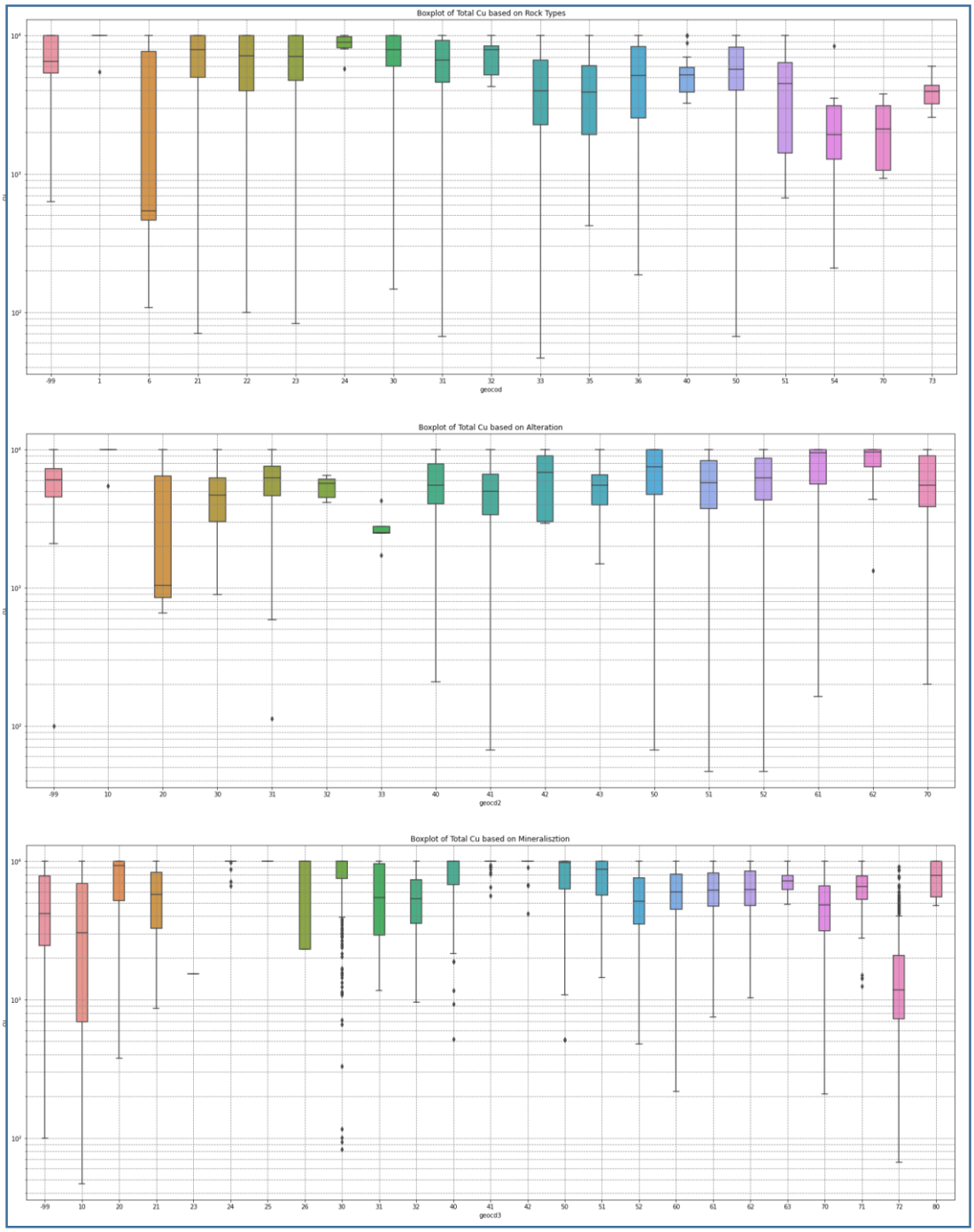


Figure 2 Copper distribution based on the three categories in different features

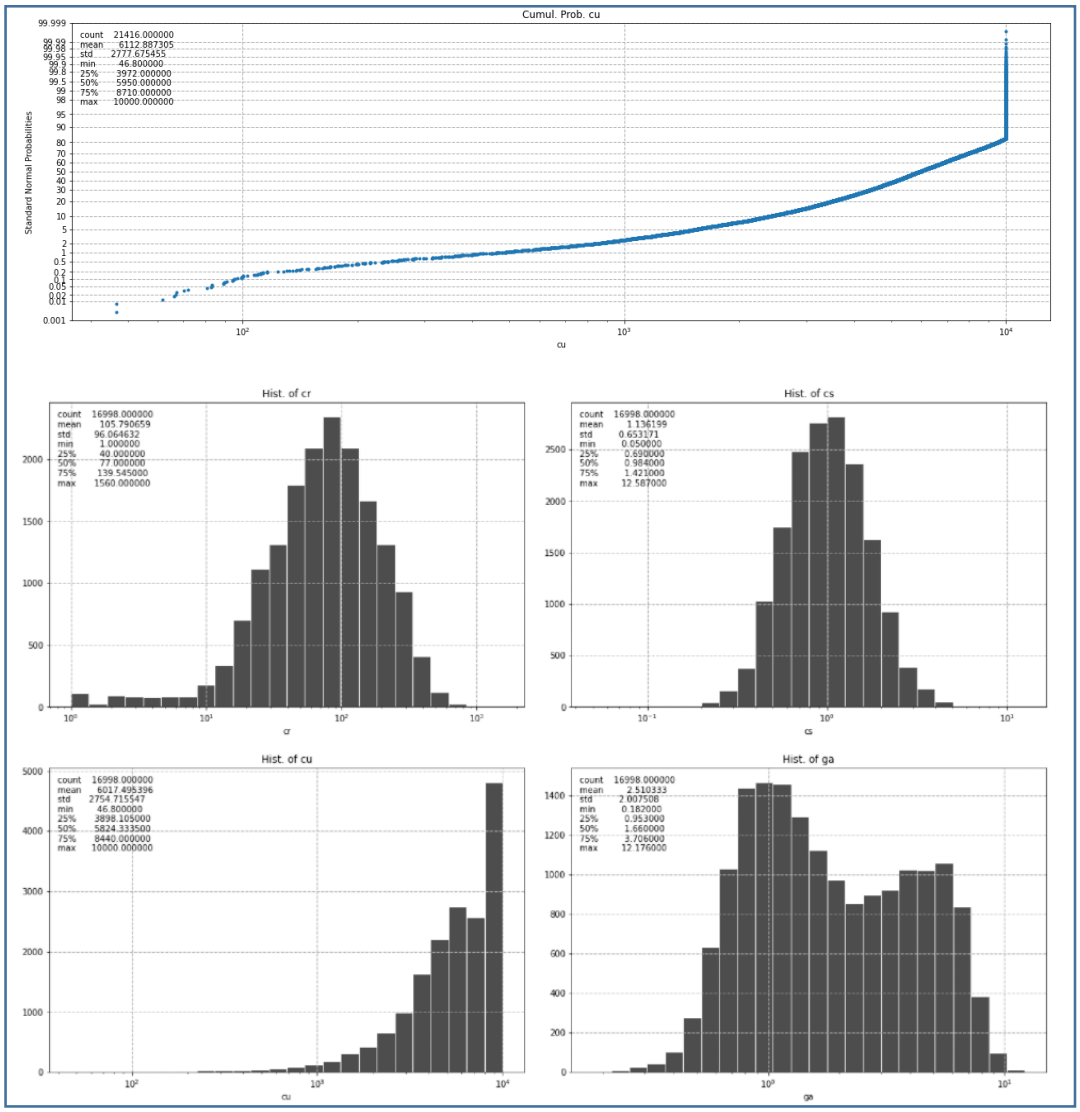


Figure 3 Probability plot of Cu values (top) and Histogram Cr, Cs, Ga and Cu distribution (below)

In the probability plot in figure 4, the variables show a clearer distinction of the populations in the alteration types, a characteristic which is not obvious in the other geological units. This suggests that the alteration is a huge factor when establishing domains and the distribution of most of the elements may vary by alteration.

Finally, the spatial distribution of the samples was also visualized in two dimensions as shown in figure 5, with preferential sampling on high-value areas, especially on its central portion, where it shows a north-east south-west trend of high values.

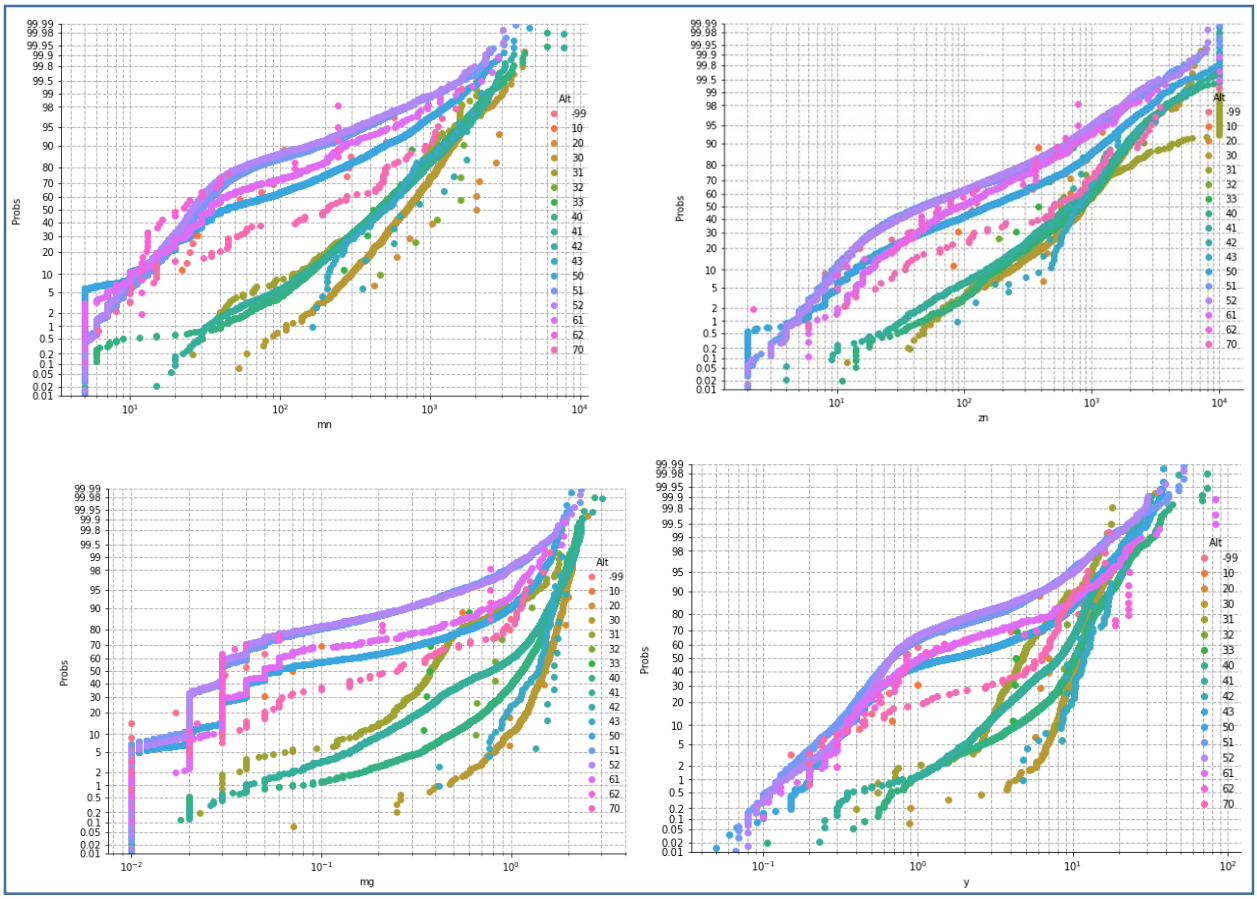


Figure 4 Distribution of Mn, Zn, Mg and Y by alteration. Two distinct populations can be observed

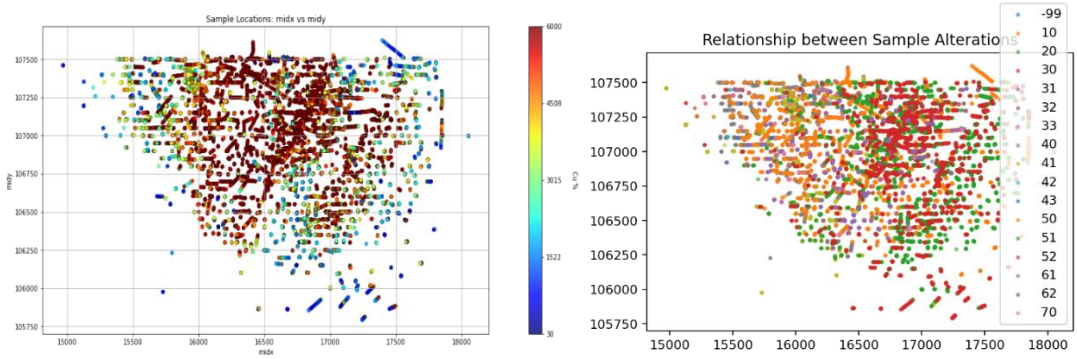


Figure 5 Location map of samples with copper concentration (left) and the various alterations units (right)

From the figures, alteration type 52 seems to be dominant alteration which host most of the high-grade copper values. Alteration type -99 is an unverified alteration and hence not regarded.

To understand the spatial correlation between elements, a correlation matrix was computed which showed poor correlation of copper with other elements in the matrix even though some trace elements of the deposit show some form of correlation with each other. Mg, Al, Ga shows a good correlation.

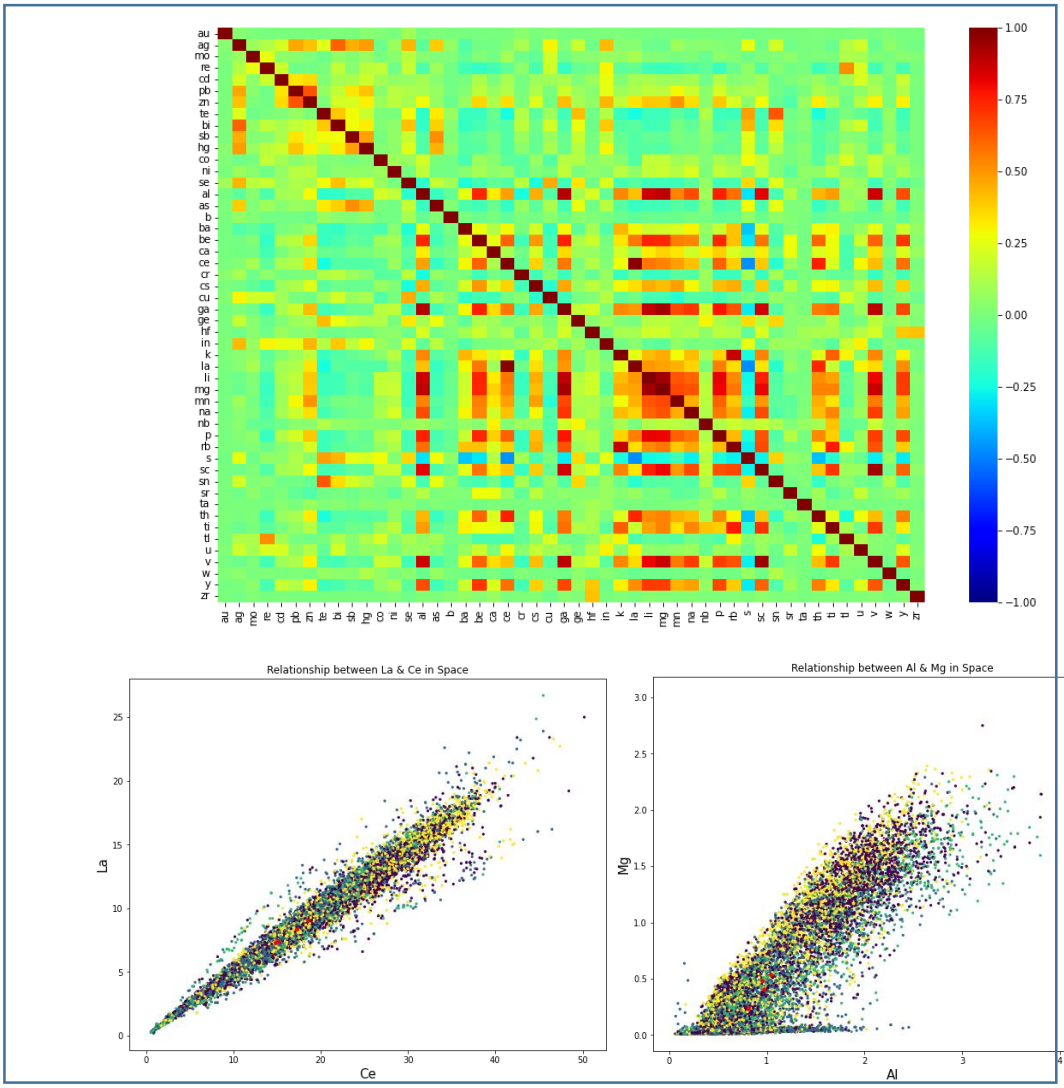


Figure 6 Correlation matrix of 50 continuous variables (top) and scatter plot showing correlation of La and Ce, and Al and Mg respectively. Color codes didn't matter at this point.

In summary, the structure of the data is well understood. Deep samples are removed and variables whose results are unaccounted for are redundant. We see how the samples are distributed in space.

In the next section, unsupervised classification using K means will be conducted. Most of the results will be shown in 3D.

3. Cluster analysis - K Means

The K-Means procedure is one of the most popular machine learning algorithms used in cluster analysis, due to its simplicity, interpretability and application to large amounts of data (Adams, 2018). It is most useful for creating a small number of clusters from many observations. Due to the large number of possible clusters that can be formed, the quality of the output is not guaranteed. The K means algorithm clusters data by separating observations into groups of equal variances, minimizing a phenomenon known as the *inertia* or within-cluster sum-of-squares as shown in the equation below (Davies & Bouldin, 1979).

Clustering was done using the web application Jupyter Notebook, with Python 3.6.5 installed via Anaconda; processor AMD Ryzen 5 3600 6-Core Processor 3.60 GHz, with 16.0GB RAM, Windows 10, 64 bit.

Two important factors that drove this analysis were the number of clusters to choose and the selection of variables to drive the clustering process. The performance of the clustering algorithm depends on the value of K. Therefore, we performed the well-known *elbow analysis* to determine the optimal number of clusters as well as a set of values for *k*. It is also important that the number of values considered should reflect the specific characteristics of the data sets which is the main motivation for performing data clustering (Pham et al., 2005).

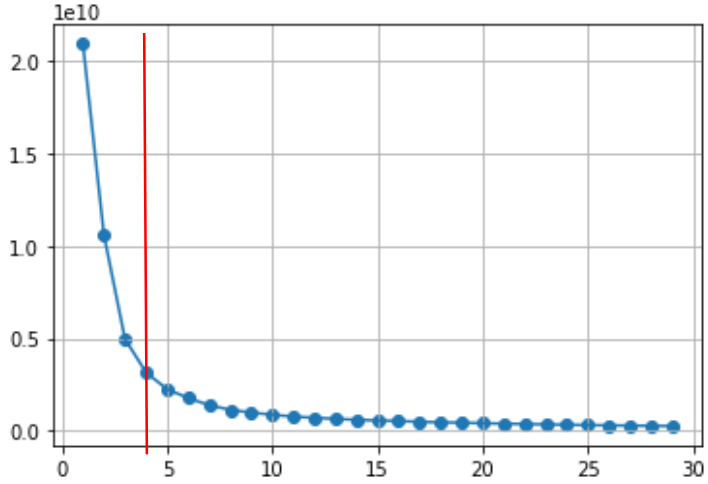


Figure 7 Determination of optimum cluster number using the elbow method. Optimum number was set at 4

The selection of variables from the geochemical data to inform the clustering process forms part of inputting domain knowledge to aid clustering since our objective is to reproduce the geology based on the geochemistry. Six elements associated with the porphyry copper deposit (Al, Ga, Mg, Li, Sc & V) were used. It is important to note that the type of domains formed is a factor of the variables selected (Faraj & Ortiz, 2021). Since most of the elements show significant changes with the alteration, we're expecting our domains to be more consistent with the alteration than the rest of the geology.

The following display shows a three-dimension clustered data of the selected variables with the elbow method optimal number of 4 as well as two other cluster number values (3 and 10).

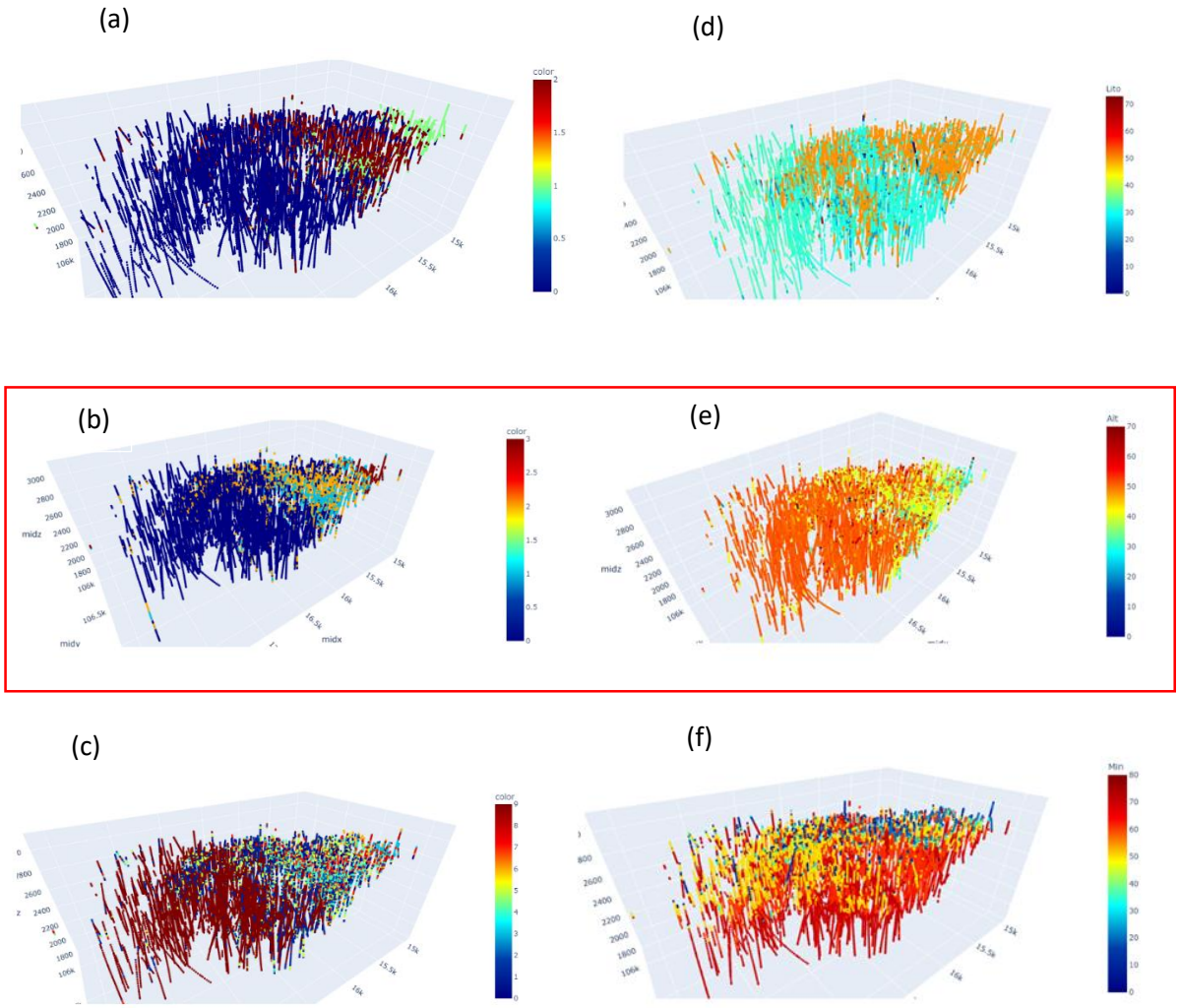


Figure 8 A 3D display of clustered drill hole data showing for cluster numbers 3, 4 and 10 represented as a, b, c respectively and the lithology, alteration and mineralization as d, e, f respectively.

It can be observed that all the three clustered data follow the alteration pattern better than either lithology or mineralization. Having established this fact, clustering with four clusters seems the better of the other cluster numbers in reproducing the alteration features. This validates the selection of the optimum number (4) by the elbow method. A cluster number of 10 was used because of its proximity to the total number of alteration features but it could not reproduce the alteration due to the larger volumes of the predominant features of the unit which overshadowed the fewer units.

Majority of the cluster labelled '0' (in blue colour) falls within alteration code 50 and 51. Clusters labelled 1 and 2 fall in alteration code 40 and 41 while the final cluster lies in the alteration 30. The K means algorithm assigns a cluster to a geological feature based on the highest number of clustered elements found in each feature. For example, if majority of the members in a particular cluster belongs to a particular lithology, K means predicts the cluster as that lithology.

4. Validation

To further validate the predicted clustered domains, a confusion matrix was computed. The confusion matrix shows the ways in which the algorithm is confused when it makes predictions, and highlights the errors made by the classifier. Since there are three geological categories present, this section validates the clustered data by these categories.

The major lithological groups present in the data were represented by codes 31 and 50 as seen in figure 9. K means prediction for the major groups attained an average accuracy of about 80% when compared to the logged lithology. The spatial distribution of the logged lithology showed that about 95% of the lithology with code 31 were rightly identified by the clustering algorithm. Rock code 50 which is the most abundant lithology was also rightly predicted. However, about 30% of lithology code 50 was misclassified as 30 by the algorithm as well as a few blocks of lithology code 30 was misclassified as 50. Due to the large number of lithologies present, misclassifications of the fewer groups are expected which leads to the decline in accuracy as it becomes trickier predicting delicate differences in units that are not largely represented or have significant similarities with the major groups.

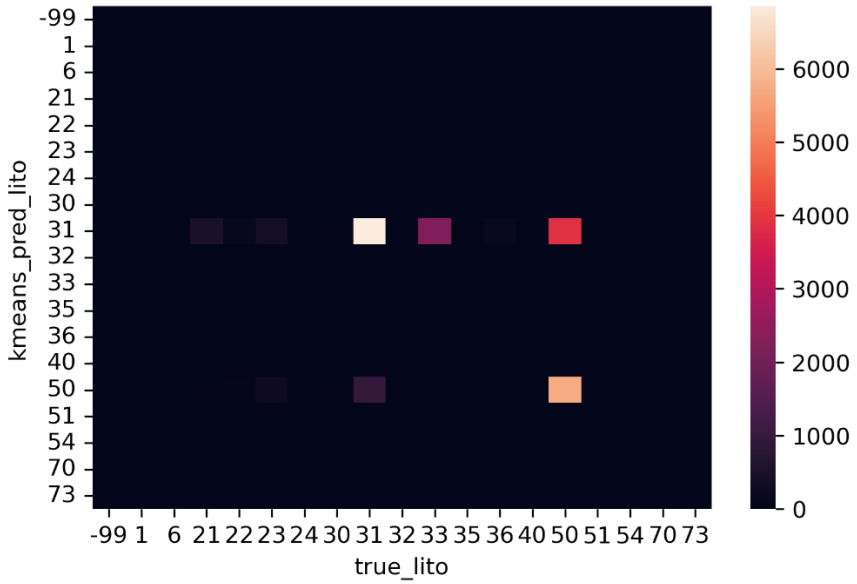


Figure 9 Confusion matrix of K means predicted units versus logged lithology

The logged alteration units contained five major units (40, 41, 50, 51, 52) with code 51 being the most abundant although the difference is not large. The algorithm predicted accurately for alteration codes 51 and 40. However, alterations that are close to code 51 are predicted as 51 as shown in figure 10. This may be due to close similarities in alterations which was evident in the grades. As captured in the probability plot in figure 5, the algorithm identifies two distinct groups of alterations. Similarly, misclassifications can be due to the insignificant representation of other alterations or due to their similarities.

The mineralogy exhibited a monopoly of predicted mineralization zones. The algorithm assigned every mineralogical zone to code 70. The algorithm performed poorly by not being able to predict the various mineralogical units present in the data. This shown in figure 11.

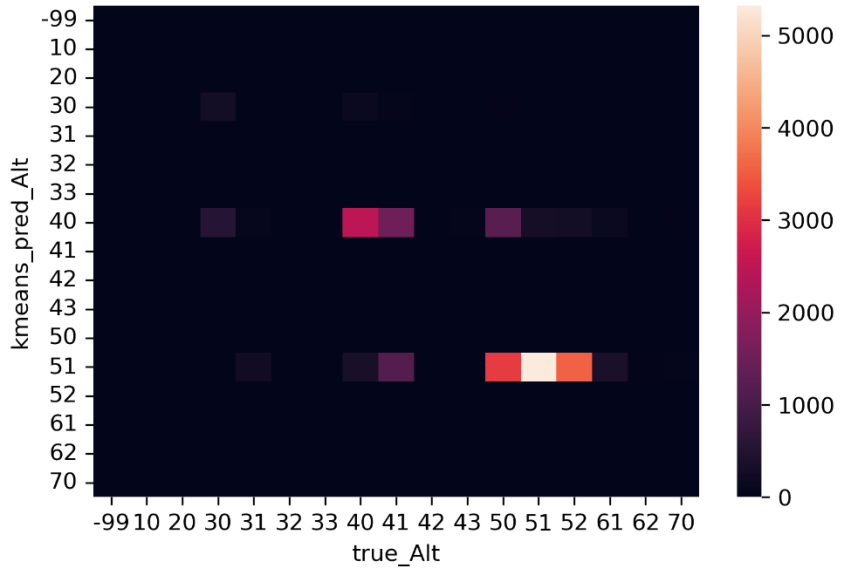


Figure 10 Confusion matrix of predicted alteration zones versus logged alteration

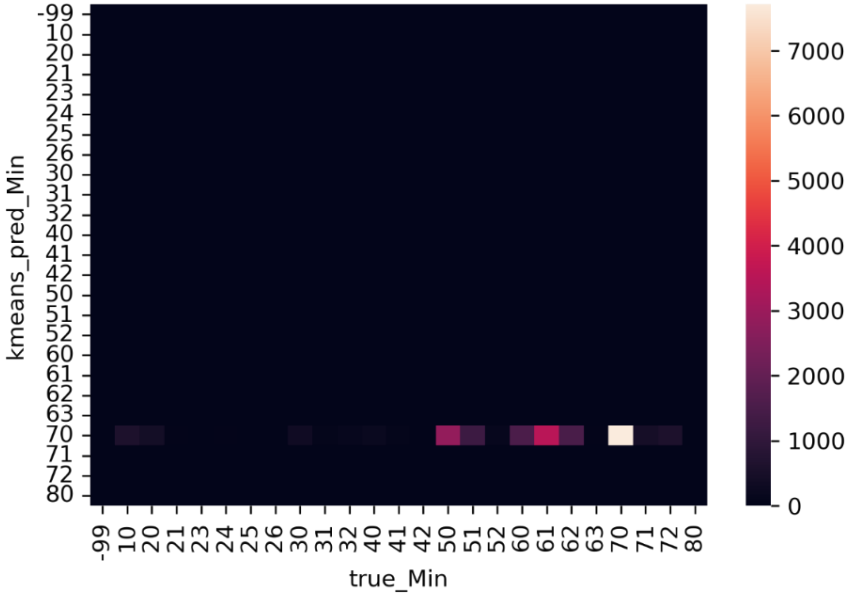


Figure 11 Confusion matrix of predicted mineralization zones versus actual mineralization

5. Conclusion

The results from this study show that although very effective and is one of the most used algorithms in machine learning, the sole application of the classical k-means is quite handicapped in geological modelling, despite the spatial contiguity it exhibits. Outliers present in the data were not properly captured which accounts for some of the misclassification.

According to the model, clusters delineated by the algorithm show some form of consistency with the logged alteration unit but forms an insignificant correlation with the rest of the categorical variables. The confusion matrix revealed the major flaws of the algorithm in its inability to classify units that are similar

or close to each other. Even though the clustered model showed a correlation with the alteration, the confusion matrix exposed its weakness in the number of misclassified units. The clustering resulted in a mixed-up population with the major units overlapping each other. The probability plot of the alteration unit indicates the presence of two major populations.

A more adequate approach is needed to account also for the geographic distribution of samples, which is done by some modern clustering techniques, such as the local autocorrelation-based clustering algorithm. The selection of variables that do not properly represent the difference in the geological units would lead to a poor discrimination by algorithm, reducing the accuracy of the model.

6. Acknowledgments

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2017-04200 and RGPAS-2017-507956.

7. References

Adams, R. P. (2018). *K-Means Clustering and Related Algorithms*. Princeton University, 18 p.

Bosch, M., Zamora, M., & Utama, W. (2002). Lithology discrimination from physical rock properties. *Geophysics*, *67*(2), 573–581. https://doi.org/10.1190/1.1468618

Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-1*(2), 224–227. https://doi.org/10.1109/TPAMI.1979.4766909

Emery, X., & Ortiz, J.M. (2005) Estimation of mineral resources using grade domains: critical analysis and a suggested methodology, Journal of the South African Institute of Mining and Metallurgy, 105(4): 247-255.

Faraj, F., & Ortiz, J. M. (2021). A Simple Unsupervised Classification Workflow for Defining Geological Domains Using Multivariate Data. *Mining, Metallurgy & Exploration*, *38*(3), 1609–1623. https://doi.org/10.1007/s42461-021-00428-5

Moreira, G. de C., Modena, R. C. C., Costa, J. F. C. L., & Marques, D. M. (2021). A workflow for defining geological domains using machine learning and geostatistics. *Tecnologia Em Metalurgia, Materiais e Mineração*, *18*, e2472. https://doi.org/10.4322/2176-1523.20212472

Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2005). Selection of *K* in *K* -means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *219*(1), 103–119. https://doi.org/10.1243/095440605X8298

Sterk, R., de Jong, K., Partington, G., Kerkvliet, S., & van de Ven, M. (2019). *Domaining in Mineral Resource Estimation: A Stock-Take of 2019 Common Practice*. Mining Geology 2019 Conference, Perth 25-26 November, 2019, 13 p.

Yasrebi, A. B., Afzal, P., Wetherelt, A., Foster, P., & Esfahanipour, R. (2013). Correlation between geology and concentration-volume fractal models: Significance for Cu and Mo mineralized zones separation in the Kahang porphyry deposit (Central Iran). *Geologica Carpathica*, *64*(2), 153–163. https://doi.org/10.2478/geoca-2013-0011

A review of grade control methods in open cast mining[1]

Noble E. Potakey (n.potakey@queensu.ca)
Julian M. Ortiz (julian.ortiz@queensu.ca)

Abstract

Ore grade control is an important part of a mine's short-term planning to define which material is considered as ore or waste. The effectiveness of grade control is subject to a compendium of factors such as sampling errors, conditional bias introduced by grade estimators, accurate definition of dig limits, and blast-induced rock movements. In this scenario, misclassification of the mined material which is mainly ascribable to the lack of absolute knowledge about real grade distribution, is our major concern. This paper reviews the state-of-the-art grade control practices used for classifying ore and waste in open pit mining. Common approaches include the classical and distance weighting estimation techniques, geostatistical methods such kriging, and simulation-based methods. Theoretical review shows that conditional simulation is a better classifier of ore and waste than estimation methods due to its ability to account for grade uncertainties and the different optimization algorithms it provides to access economic consequences of ore/waste decisions. Subsequently, a novel machine learning approach using Elliptical Radial Basis Function Network (ERBFN) and Support Vector Regression (SVR) was discussed. Results from a case study show that SVR achieved an 8%, 1.12% and 1.16% reduction in misclassified material relative to inverse distance, ordinary kriging, and simulation respectively. The ERBFN model also obtained a decrease in misclassified material of 12%, 5.4% and 5.7% compared to inverse distance, ordinary kriging, and simulation-based approaches, respectively.

1. Introduction

Considering the extensive nature of mining operations which involves making decisions regarding large volumes of materials taking place over limited period of times, blending of waste with ore and ore with waste is inevitable. Nevertheless, geologists and mine engineers must ensure that this situation is reduced to the barest minimum during excavation, and that ore and waste are differentiated. The decision of which material is ore or waste is very crucial for the mine's profitability since it is the last opportunity for the mining company to achieve its estimated revenue, and errors at this stage are very costly and difficult to reverse due to its proximity to the production stage (Rossi & Deutsch, 2014). To help make better decisions in classifying ore and waste, and to select the destination of each parcel of material mined, mines perform grade control.

Ore grade control is a compendium of procedures and practices usually involving blast hole sampling, grade estimation, ore/waste classification, blast-induced rock movement measurements, defining dig limits among others. It is done to identify which material is ore or waste and to ensure that the mill is fed

[1] Cite as: Potakey N. E., Ortiz J. M. (2022) A review of grade control methods in open cast mining, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-04, 59-70.

with the right grade of material. In this scenario, material classification based on the grade assignment is discussed. A volume of material is classified as ore or waste based on the grade assigned to it from grade prediction methods and by comparison to a cut-off grade (Abzalov, 2016). When the grade assigned to a material falls below a given cut-off grade, it is considered as waste and sent to the waste dump while material with grade above the cut-off grade is sent to the mill. However, when the assigned grades are not consistent with the actual grade distribution, it results in a misclassification, and materials are sent to the wrong destination. The figure below shows the basic issue of misclassification where a scatterplot of true grades for each block are plotted against the corresponding predicted grades.

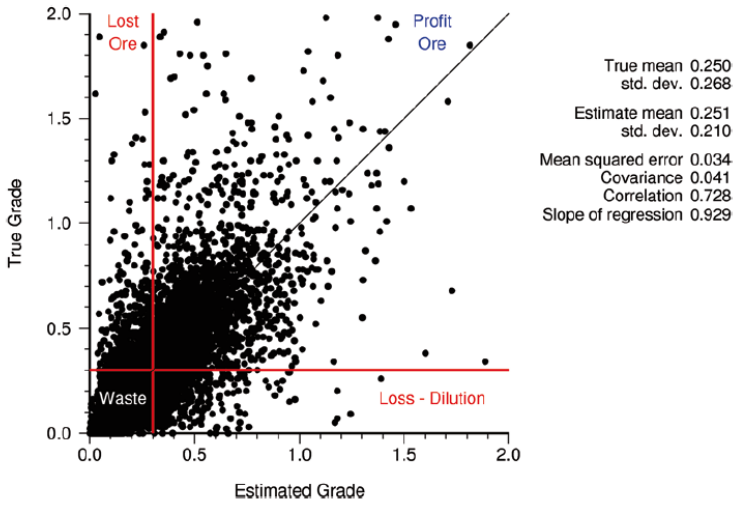


Figure 1: Misclassification in grade control. Source: (Rossi & Deutsch, 2014).

Common approaches include the classical nearest neighbour method where grades of the closest blasthole sample are assigned directly to a block model, and the inverse distance weighting estimation method which assign grades by calculating the weight of each sample that is inversely proportional to the distance of the estimation location (Ortiz, 2020). Geostatistical methods such as kriging were introduced in the early 1950s. Kriging minimizes the estimation variance under certain conditions (Rossi & Deutsch, 2014). Conditional simulation methods have been the most advocated technique in recent years to predict grades and their uncertainty, thus addressing some of the shortcomings of kriging and the classical methods. This method simulates grades at given locations and can be combined with different optimization algorithms where the material types (ore or waste) are evaluated against all simulated realizations and the optimum destination or material type at each location determined (Vasylchuk & Deutsch, 2018).

In this paper (Vasylchuk & Deutsch, 2018), the authors review four grade interpolation methods:

1. Nearest neighbour polygonal method;
2. Inverse distance weighting method;
3. Ordinary kriging; and
4. Simulation-based methods.

The paper also discusses the performance of two novel machine learning (ML) algorithms that have been used recently in grade control: Elliptical Radial Basis Function Network (ERBFN) and Support Vector Regression (SVR). The results from these methods are analysed with the help of the case studies presented to show which method is most efficient for grade control.

In the next sections, grade control methods are reviewed and critically assessed in a simple theoretical framework. Results obtained in the literature from numerical experiments conducted to compare the effectiveness of simulation versus different estimation methods, as well as results on a comparative case study at the Carmen de Andacollo copper mine (Chile) are presented. Summary and conclusions follow.

2. Methods for grade prediction

Grade prediction is probably the most crucial aspect of grade control because it forms the basis for selecting ore and waste zones. The main purpose of estimation is to predict the grade of the variable at unsampled locations in the block model which is the premise for classification. Samples taken from blast holes are analyzed from which a quantitative model of the ore body is constructed by interpolating and extrapolating between these samples to account for the grade in areas that were not sampled. This is with the assumption that all the sample locations and the unsampled location belong to the same domain. There are various methods developed for performing grade prediction in grade control, however this paper discusses the most common ones in the industry.

2.1. Nearest Neighbour (NN) estimation method

The nearest neighbor method is one of the simplest approaches to grade estimation. As a variant of the polygonal method, the nearest neighbor assigns the grade of the closest blast hole sample to the entire unsampled block (Vasylchuk, 2019). Since the weight of each sample is an important factor in estimating, NN method determines the weight of the samples by assigning all the weight to the closest sample and every other sample gets a weight of zero. In relation to the cut-off grade, the NN method regards the estimation location as ore if the grade of the closest sample is larger than the cut-off grade. The value at the unsampled location is then calculated as shown in the equation below:

$$z^*_{NN}(u_o) = \lambda_o^{NN} + \sum_{i=1}^{n} \lambda_i^{NN} z(u_i) \qquad (1)$$

$$\lambda_i^{NN} = \begin{cases} 1 & if\ u_i\ is\ closest\ to\ u_o \\ 0 & otherwise \end{cases} \qquad i = 1, \dots, n$$

$$\lambda_0^{NN} = 0$$

where $z^*_{NN}(U_o)$ is the nearest neighbor estimate at the unsampled location (u_o), λ^{NN} are the nearest neighbor weights of the samples, and $z(u_i)$ is the known grade of the samples for which $i = 1, \dots, n$ are the total number of samples.

One main strength of the nearest neighbour method is that it does not smooth estimated values (Kapageridis, 2014). However, study has shown that this method is not particularly accurate and cannot be trusted because it neither takes into account the spatial continuity of the grade nor the redundancy in the information (Ortiz, 2020). The discrepancies in this method are known to be larger that other estimators, and for many deposits that have positively skewed distributions, significant errors in the estimate occur in the individual blocks leading to proclivity to overestimate the average grade and

underestimate tonnage above cut-off (Rossi & Deutsch, 2014). Nonetheless the nearest neighbour method can be used as a checking tool.

2.2. Inverse Distance Weighting (IDW) estimation method

The inverse distance weighting technique is an enhancement of the classical polygonal method, and is most suitable for uniform orebodies (Abuntori et al., 2021). This method is used to estimate grade values using several nearby blasthole sample grades to obtain a weighted average for each block as shown in Figure 2. The calculation of the estimate is shown below:

$$z^*_{IDW}(u_o) = \lambda_0^{IDW} + \sum_{i=1}^{n} \lambda_i^{IDW} z(u_i) \qquad i = 1, \dots, n \qquad (2)$$

where $z^*_{IDW}(u_o)$ is the inverse distance estimate at the unsampled location (u_o), λ^{IDW} is the inverse distance weight assigned to each known sample and $z(u_i)$ is the grade of each known sample for which $i = 0, \dots, n$ are the total number of samples. In this scenario, the weights assigned to each sample are inversely proportional to the distance from estimation location and are calculated as shown in equation 3 below. Each sample is weighted based on its proximity to the location to be estimated.

$$\lambda_i^{IDW} = \frac{{}^{1}/_{(c+d_{io}^{w})}}{\sum_{i=1}^{n} {}^{1}/_{(c+d_{io}^{w})}} \qquad i = 1, \dots, n \qquad (3)$$

$$\lambda_0^{IDW} = 0$$

for which λ_i^{IDW} is the inverse distance weight of sample i, d_{io} is the distance between the estimation location and sample i, w is the inverse distance weighting power and c is a small constant for numerical stability or computational reasons. When the weighting power approaches zero, the weights become similar and is calculated as the arithmetic average of the samples (Abzalov, 2016). On the other hand, a larger weighting power assigns all the weight to the closest sample making the inverse distance weight similar to the result of the polygonal nearest neighbour method (Ortiz, 2020). In practice, the most frequently used weighting power is 2, however powers of 1 and 3 are also used for estimation. Even though the inverse distance weighting method provides better estimates than the nearest neighbor method, it does not account for the details of the data configuration or the varying anisotropy at different scales (Vasylchuk, 2019).

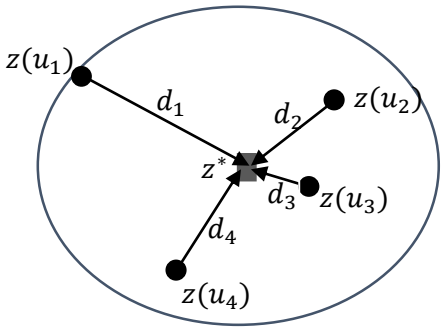


Figure 2: Inverse Distance Weighting method.

2.3. Kriging-based estimation method

Geostatistics, developed in the early 1950s, forms the basis for kriging (Krige, 1951). The geostatistical concept provides the platform for describing and modelling the spatial continuities of the regionalised variables (in this case the grade values) and allows incorporation of the continuity factors into the regression techniques used for the spatial predictions (Abzalov, 2016). Kriging is a collection of generalized linear regression techniques based on calculating optimal weights that minimize the expected error variance or the estimation variance (Ortiz, 2020). It produces an estimate that is a weighted linear combination of the data, minimizes the estimation error, hence, it called the Best Linear Unbiased Estimator (BLUE).

Kriging-based grade control came to light in open pit mines during the 1980s (Deutsch et al., 2000). Different types of kriging algorithms have been used in grade control, but most commonly ordinary kriging (OK) particularly in gold mines in Northern Nevada (Rossi & Deutsch, 2014). Other types of kriging have also been applied such as the indicator kriging and not too popular fuzzy kriging (González, 2012). Although not very common, simple kriging (SK) has also been used to estimate but more often used as a checking tool.

Ordinary kriging

Ordinary kriging (OK) is a robust estimator which assumes that the local mean is unknown (unlike SK which assumes a known mean), but constant within the estimation neighborhood based on the quasi second order stationarity assumption (Ortiz, 2020). To guarantee global unbiasedness, OK constrains the sum of the weights to be 1.0, and as a result the mean does not need to be known (Abuntori et al., 2021). The weights used in kriging directly depend on the choice of a variogram model for the data set. The variogram model, which is a prerequisite, enables the kriging algorithm to obtain insight on the anisotropy in the grade distribution. The ordinary kriging estimate is summarized in the equation below.

$$Ordinary\ kriging\ estimator,\ Z^*_{OK}(u_o) = \sum_{i=1}^{n} \lambda_i^{OK} Z(u_i) \qquad (4)$$

where $Z^*_{OK}(u_o)$ is the ordinary kriging value at the unsampled location (u_o), λ_i^{OK} the weight assigned to each known sample (u_i), and $Z(u_i)$ is the grade of each known sample i for which $i = 1, \dots, n$ are the total number of samples. The kriging variance which measures the quality of the estimation is given by

$$Ordinary\ kriging\ variance,\ \sigma^2_{OK}(u_o) = \sigma_0^2 - \sum_{i=1}^{n} \lambda_i^{OK} C_{io} - \mu \qquad (5)$$

where $\sigma^2_{OK}(u_o)$ is the kriging variance at the estimation location (u_o), σ_0^2 is the variance of the distribution, λ_i^{OK} is the weight assigned to the known sample i, C_{io} is the covariance between sample i and the estimation location (u_o), and μ is the Lagrange multiplier which is an additional parameter to help in optimality.

There are other variants of the kriging method which have proven successful in many operations such as the Breakeven Indicator Method (BEI) (Vasylchuk, 2016). The BEI grade control method is a blend of both indicator and grade kriging. It uses an ore/waste indicator variable to predict the probability of ore occurrence at a given location (Rossi & Deutsch, 2014). The indicator variable is then used to define ore or waste probability of the estimated value based on the grade of the blast holes, and the expected revenue is determined. This method was used in copper-molybdenum Ujina open-pit in Chile, together with and the classical inverse distance weighting method and the results were compared to a reference

model. The BEI showed a relatively better performance than IDW and a summary of the results is shown in subsequent sections.

In summary, kriging provides good frameworks for predicting grades that are locally accurate estimates, however, the premise of estimating based on the minimization of estimation variance is not optimal for grade control (Srivastava, 1987). (Rossi & Deutsch, 2014) recounts that kriging has been only slightly more successful at grade control compared to the other classical methods because of the inherent smoothing and the inability to quantify the spatial uncertainty as shown in Figure 3.

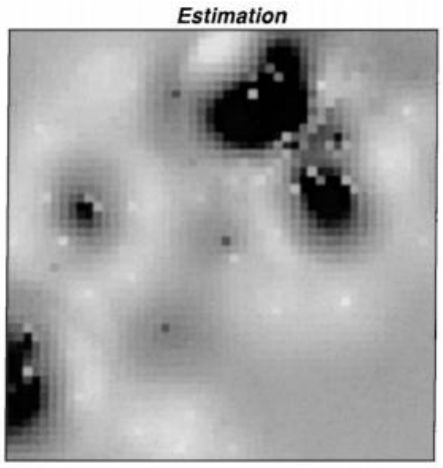


Figure 3: Smoothing effect of kriging.

2.4. Simulation-based method

The use of simulation as a predicting tool has been a cutting-edge method in grade control. The commonest simulation method used to model the realistic variability of a deposit is the Sequential Gaussian Simulation (SGS) (Vasylchuk & Deutsch, 2018). This method assigns grade to blocks and also takes into account the uncertainty in the grade distribution that can be later used for assessing economic consequences of grade control decisions, a feature lacked by the traditional estimation methods (Vasylchuk, 2019). The grade assignment process involves a series of steps such as data declustering, normal score transformation of declustered data, simulating a value from the conditional distribution and back-transforming simulated values. The simulation process returns a range of probable values from a conditional cumulative distribution function (cdf) as shown in the equation below:

$$F(u; z|(n)) = Prob\{Z(u) \leq z|(n)\} \quad (6)$$

where $F(u; z|(n))$ is the cumulative frequency distribution curve, $Z(u)$ accounts for the uncertainty in the unknown true value, and (n) represents local conditioning blast hole data within the specific neighborhood of location (u).

In order to evaluate which realization will produce the optimum classification of materials, taking into consideration its economic impact on the operation, simulation-based methods incorporate certain optimization algorithms such as minimum loss and maximum profit functions (Deutsch et al., 2000; Dimitrakopoulos & Godoy, 2014; Vasylchuk, 2016). Moreover, the economic consequences of sending ore to the waste dump is different from sending waste to the mill, hence an optimal selection criterion is needed to account for these asymmetric economic impacts through the optimization algorithms that

simulation provides. The application of the economic classification functions for grade control does not only account for the penalties of each decision, but also provide essential information for non-linear metal recoveries or any other geo-metallurgical attribute of interest by adopting the economic functions (Wambeke & Benndorf, 2017).

The minimum expected loss method consists of computing the expected loss associated with each classification and selecting the classification for which the expected loss is minimal. Several mathematical expressions have been provided by different authors for the 'minimum loss' function, but for the purposes of subsequent comparison, a simplified version of the loss function by (Vasylchuk, 2016) is presented, as shown in equation 7.

$$The\ loss\ function, g(u;z,z_c) = \begin{cases} 0,\ for\ a\ correct\ decision \\ (z(u)-z_c)\times b_1, for\ an\ incorrect\ waste\ decision \\ (z_c - z(u))\times b_2, for\ an\ incorrect\ ore\ decision \end{cases}$$

Hence, $the\ Expected\ loss\ decision = E[g(u;Z,z_c)]$ (7)

where $g(u;Z,z_c)$ represents the loss function, $E[g(u;Z,z_c)]$ is the expected loss over multiple realizations at location (u), z_c is the cut-off grade, $z(u)$ is the simulated grade value at location (u), b_1 and b_2 are the penalty coefficients for underestimating and overestimating respectively. For instance, a block will be selected as ore if the expected loss for mining the block as ore is less than the expected loss for mining it as waste.

Similarly, for maximum profit function, a block will be selected as ore if the expected profit for mining the block as ore exceeds the expected profit for mining it as waste, and vice versa. More information on concept of minimum expected loss or maximum profit as a basis for classification decisions can be found in (Rossi & Deutsch, 2014; Vasylchuk, 2016; Verly, 2005; Wambeke & Benndorf, 2017).

2.5. Machine Learning (ML) method

Technological advancements in recent years have enabled computers to process large amounts of information within the shortest possible time. Machine learning algorithms (MLA) are a collection of advanced statistical methods empowered by high-level computers to provide flexibility and simplicity when integrating and recognizing complicated patterns in data, which is a difficult task with linear geostatistical workflows. Machine learning is already incorporated to solve Earth Sciences problems (Deutsch et al., 2016) but is not fully integrated into the geometallurgical workflows that model and optimize the processes leading to the extraction and recovery of minerals and metals (Ortiz, 2019). As far as this review is concerned, the application of machine learning methods for grade assignment and classification in grade control is a novel enterprise with not too many applications. ML algorithms such as Artificial Neural Networks have been used for mineral resource estimation (Abuntori et al., 2021) but have not been explored in grade control. The application of ML is to enhance the accuracy of predicted grade values, and to make better decisions concerning the classification of mined material.

A recent work by (Da Silva et al., 2020) demonstrated the application of ensemble ML methods for grade control in the Carmen de Andacollo copper mine in Chile. Two algorithms were used in their work: Elliptical Radial Basis Function Network (ERBFN) and Support Vector Regression (SVR). The two ML algorithms are trained which was preceded by the tuning of their respective hyperparameters. For instance, the number of nodes that constitute a hidden layer in the network was defined for ERBFN. Once

the nodes are defined, ERBFN algorithm trains the data and assigns a new networking system to each node to make grade predictions. The different predictions obtained are averaged to form a trend model which is further used as a secondary variable to assess the consequences of each grade control decision on an intrinsic collocated cokriging framework. Just as kriging, variogram models were generated and search parameters defined.

In the case of SVR, the data is divided into a training, validation, and test dataset. Different reference models are defined from which predictions are generated for the validation set. The predicted values and the blasthole values are then fed into a trained meta model from these pairs of values, and then final grade predictions are made. Just like ERBFN, the SVR algorithm optimizes the best scenario for ore and waste by using the final predicted model in a collocated cokriging framework. Finally, based on a break-even cut-off grade defined by the mine operations, the destination of each material is determined. Details of this work can be found in (Da Silva et al., 2020).

3. Results and Discussions

The results presented here are case studies reviewed from (Da Silva et al., 2020; Rossi & Deutsch, 2014; Vasylchuk, 2016) which demonstrate the performance of grade prediction methods. The first scenario is a study in a copper-molybdenum Ujina open pit mine in Northern Chile where the outcomes of the Break-even Indicator (BEI) method, and the inverse distance weighting (IDW) method are compared to a reference model in terms of their ability to classify and provide destination for materials. The results show that BEI is superior to IDW in that it produced results closer to the reference model. The study further stated that the simulation based approached used here produced results similar yet slightly better than the BEI. Meanwhile, Ordinary kriging (OK) produced a marginally inferior result. Only results for BEI and IDW are presented here with respect to the reference model for tonnages and total Cu grade for different destinations. Proximity of a value to 1.0 indicates better performance of the method. A factor greater than 1 implies overestimation with respect to the reference model. Details of this work can be found in (Rossi & Deutsch, 2014).

Table 1: Performance of IDW and BEI models with respect to an SGS reference model. Source: (Rossi & Deutsch, 2014).

Destination code	**Tonnage w.r.t reference**		**TCu Grade w.r.t reference**	
	IDW	BEI	IDW	BEI
SAL	1.10	1.10	0.91	0.92
SME	1.16	1.09	1.06	1.00
SBA	0.18	0.45	1.15	1.01
SMR	0.50	0.43	1.36	1.01
SAS	0.55	0.87	1.02	0.95
OXA	1.29	1.13	0.85	0.93
OXB	1.16	1.98	1.08	0.98
OXL	0.44	1.49	1.54	1.41
MIX	0.52	0.71	0.90	0.78
TOTAL	1.16	1.11	0.84	0.89

In the second scenario, a numerical experiment was conducted to evaluate the effectiveness of simulation versus different estimation methods based on the losses incurred as derived from the expected loss

function. The penalty coefficients for underestimating and overestimating, b_1 and b_2 respectively are given as a ratio. A ratio of 1:1 means the consequences of under or overestimating are equal. If the ratio is asymmetric (example 1:2), it means that the penalty for overestimating is higher. Only a few ratios are shown here for illustrative purposes. The results show that the simulation-based method incurred the least losses than the other methods both on average and on individual penalty considerations.

Table 2: Incurred losses of grade control methods. Source: (Vasylchuk & Deutsch, 2018).

GC	b_1: b_2							
Methods	2.65:1	2:1	1.3:1	1:1	1:1.3	1:2	1:2.65	**Average**
NN	1208.4	433.3	257.0	206.8	246.6	387.3	1006.0	454.0
ID	900.6	333.3	204.2	166.4	203.9	331.8	894.2	369.3
OK	877.7	323.7	197.6	160.6	196.4	318.5	855.1	359.8
SK	861.9	316.5	192.4	156.0	190.3	307.2	820.9	349.8
Simulation	338.7	215.6	186.7	155.8	186.3	249.7	355.0	224.7

The last scenario involves the integration of machine learning techniques into grade control. The two ML algorithms used (ERBFN and SVR) were compared to other commonly used methods in the industry with respect to how their grade assignment led to the classification of the materials. The methods were Inverse distance (ID), ordinary kriging (OK), and an intelligent grade control (IGC) based on multivariate simulation. From the study, it was observed that ERBFN and SVR, in conjunction with the collocated cokriging framework outperformed traditional estimation methods and the simulation-based method. The efficient grade prediction process led to better classification and better prediction of material destinations. In Figure 4 below, a five-fold mean square error (MSER) validation is used to measure the performance of each model over ten blast holes considered. The red points represent the overall MSER for each method. From the study, it was observed that the inverse distance estimation (ID) obtained the highest MSER of 0.00888, while that of ERBFN and SVR obtained an MSER of 0.007 and 0.0075 respectively.

Finally, the extent of misclassification by the methods were evaluated and the study revealed that ERBFN with collocated cokriging (CCok) obtained the best performance in reducing misclassification. It obtained a reduction in misclassified material of 12% when compared to ID; 5.4% less than OK, 5.7% relative to IGC and 4.3% relative to CCok with SVR. Even though ERBFN-CCok obtained better results that SVR-CCok, the latter also obtained a reduction of 8% in misclassified material relative to ID and 1.12% and 1.16% to OK and IGC respectively as shown in Figure 5.

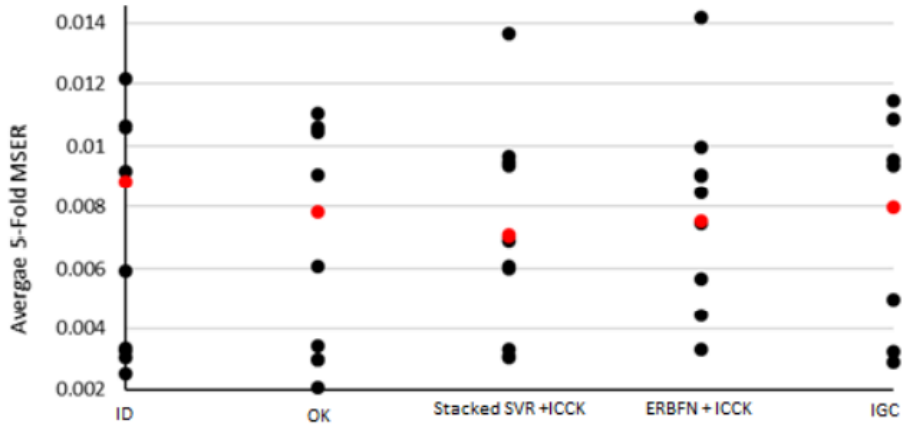


Figure 4: Mean squared error obtained from a 5-fold cross validation for each method applied (source: (Da Silva et al., 2020).

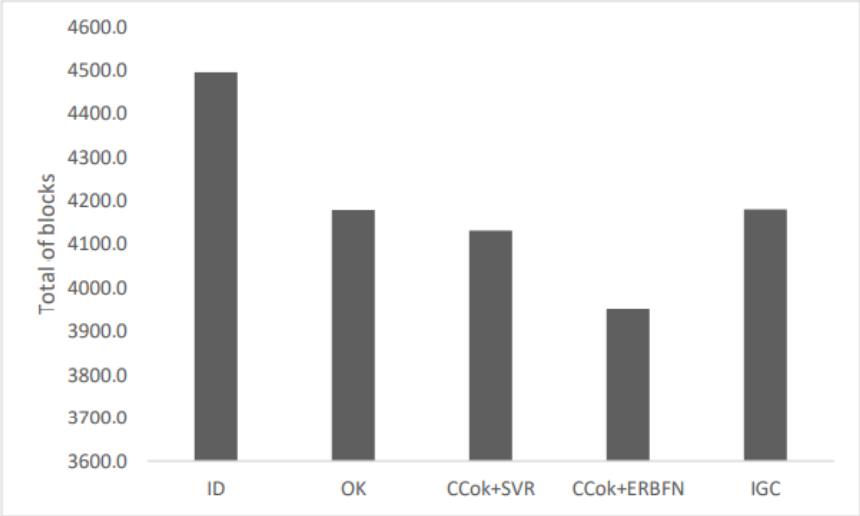


Figure 5: Total number of misclassified blocks recorded by the grade control methods (source: (Da Silva et al., 2020).

4. Conclusions

Grade control methods for classifying materials have evolved. The methods reviewed in this paper have shown that classical estimation methods can no longer be relied on in making grade control decisions. The smoothing effect and other limitations of kriging methods are a huge concern for grade control and hence do not make it an optimal method. Simulation based methods have provided good results which still make them a cutting-edge tool in grade control and is still used by most mining companies today. However, machine learning methods may become the future of grade control due to its high performance in grade prediction and significantly reducing misclassification, which is the goal of grade control.

5. Acknowledgments

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2017-04200 and RGPAS-2017-507956.

6. References

Abuntori, C. A., Al-Hassan, S., & Mireku-Gyimah, D. (2021). Assessment of Ore Grade Estimation Methods for Structurally Controlled Vein Deposits—A Review. *Ghana Mining Journal*, *21*(1), 31–44. https://doi.org/10.4314/gm.v21i1.4

Abzalov, M. (2016). *Applied Mining Geology* (Vol. 12). Springer International Publishing. https://doi.org/10.1007/978-3-319-39264-6

Da Silva, C. Z., Nisenson, J., & Boisvert, J. (2020). *Grade control with Ensembled ML: A comparative case study at the Carmen de Andacollo copper mine* [Annual report].

Deutsch, C. V., Magri, E., & Norrena, K. (2000). Optimal Grade Control Using Geostatistics and Economics. *Society for Mining, Metallurgy, and Exploraction Inc.*, *308*.

Deutsch, J. L., Palmer, K., Deutsch, C. V., Szymanski, J., & Etsell, T. H. (2016). Spatial Modeling of Geometallurgical Properties: Techniques and a Case Study. *Natural Resources Research*, *25*(2), 161–181. https://doi.org/10.1007/s11053-015-9276-x

Dimitrakopoulos, R., & Godoy, M. (2014). Grade control based on economic ore/waste classification functions and stochastic simulations: Examples, comparisons and applications. *Mining Technology*, *123*(2), 90–106. https://doi.org/10.1179/1743286314Y.0000000062

Giraldo, R., Herrera, L., & Leiva, V. (2020). Cokriging Prediction Using as Secondary Variable a Functional Random Field with Application in Environmental Pollution. *Mathematics*, *8*(8), 1305. https://doi.org/10.3390/math8081305

Giraldo, R., Mateu, J., & Delicado, P. (2017). *Cokriging and multivariate kriging methods based on data of a functional random field*. *10*(2), 30.

Gräler, B. (n.d.). *Cokriging and indicator kriging*. 21.

Kapageridis, I. K. (2014). *APPLICATION OF ARTIFICIAL NEURAL NETWORK SYSTEMS TO GRADE ESTIMATION FROM EXPLORATION DATA*. 268.

Krige, D. G. (1951). A Statistical Approaches to Some Basic Mine Valuation Problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, *52*, 119–139.

Moreira, G. de C., Modena, R. C. C., Costa, J. F. C. L., & Marques, D. M. (2021). A workflow for defining geological domains using machine learning and geostatistics. *Tecnologia Em Metalurgia, Materiais e Mineração*, *18*, e2472. https://doi.org/10.4322/2176-1523.20212472

Ortiz, J. M. (2019). Geometallurgical modeling framework, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2019, paper 2019-01, 6-16.

Ortiz, J. M. (2020). Advanced Geostatistics—Estimation. *Queen's University*, 29.

Rossi, M. E., & Deutsch, C. V. (2014). *Mineral Resource Estimation*. Springer Netherlands. https://doi.org/10.1007/978-1-4020-5717-5

Srivastava, R. M. (1987). Minimum variance or maximum profitability. *CIMM*, *80(901)*, 63–68.

Vasylchuk, Y. V. (2016). *Integrated System for Improved Grade Control in Open Pit Mines*. Master of Science in Mining Engineering thesis, University of Alberta, 150.

Vasylchuk, Y. V. (2019). *Advanced Grade Control with Multivariate Geostatistics, Blast Movement Modeling, and Optimized Dig Limits*. Doctor of Philosophy thesis, University of Alberta, 236.

Vasylchuk, Y. V., & Deutsch, C. V. (2018). Improved grade control in open pit mines. *Mining Technology*, *127*(2), 84–91. https://doi.org/10.1080/14749009.2017.1363991

Verly, G. (2005). Grade Control Classification of Ore and Waste: A Critical Review of Estimation and Simulation Based Procedures. *Mathematical Geology*, *37*(5), 451–475. https://doi.org/10.1007/s11004-005-6660-9

Wambeke, T., & Benndorf, J. (2017). A Simulation-Based Geostatistical Approach to Real-Time Reconciliation of the Grade Control Model. *Mathematical Geosciences*, *49*(1), 1–37. https://doi.org/10.1007/s11004-016-9658-6

Sampling error and its effect on grade control profit[1]

Kwame A. Ntiri (21kan4@queensu.ca)
Julian M. Ortiz (julian.ortiz@queensu.ca)

Abstract

Mineral grade prediction is a critical phase in mineral exploration and resource estimation, and it plays a vital role in the economic evaluation of mining projects. In mining, grade control is the process of identifying where the mined material will end up. Misclassification of ore grades costs money, hence the requirement for representative sample methodologies in open-pit mining is becoming increasingly critical in all mining industries. The impact of Ordinary Kriging estimation is discussed in this study. Ordinary kriging is regarded as a highly dependable method and is commonly used for estimation. A mine's blasthole data is examined. Because errors are likely to occur during grade control or even actual mining operations, data analysis is performed by introducing errors in the data. An economic analysis is performed on the various errors introduced and how they will affect the mine's profit.

1. Introduction

Mining companies want to increase the returns on the investment made in the mining operation, hence the need to optimize grade control to minimize misclassification of ore and waste. Grade control is a technique that offers selectivity for the extraction of different types of ore and waste with the aim of increasing profit or minimizing loss in the mining operation (Verly, 2005). According to the characteristics of the ore, the detailed data gathered at the grade control stage is used to separate ore from waste, a process known as "ore-waste classification," and to determine the final destination of the various material types (Abzalov et al., 2010).

The quality and quantity of the samples used determine the effectiveness of ore grade control at active mines. The simplest implementation of grade control consists of manually designing ore-waste boundaries or ore blocks on a map of blast hole grade values (Verly, 2005). In grade control, data can be viewed as a distribution, normally as a histogram, or as spatial continuity, where data is analyzed as a variogram. The main argument for using simulations is that smoothed maps obtained by kriging do not account for the uncertainty in the grade estimation or for the economic consequence of misclassification. As a result, optimum classification cannot be achieved (Verly, 2005). Geostatistical simulation methods for grade control have been used to solve optimization problems such as minimizing loss functions or maximizing the expected profits (Glacken, 1997; Deutsch, Magri and Norrena, 1999).

This paper aims to compare the estimated grade using the Nearest Neighbor, Inverse Distance Weighted, Inverse Distance Weighted Squared, and Ordinary Kriging to the true grade of a dense grid, which is the grade produced through simulation, and discuss its effect on profit. An open pit mine in Chile will be used

[1] Cite as: Ntiri K. A., Ortiz J. M. (2022) Sampling error and its effect on grade control profit, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-05, 71-84.

as a case study, with the estimated grade compared to the true grade using a cutoff. GSLIB software will be used for generating variograms and simulating the blasthole data.

2. Literature review

Grade control procedures depend on both the quality and quantity of samples. Just improving the sample quality does not always lead to better defined ore and waste blocks if the spacing chosen is too broad. The primary goal of grade control in mines is to distinguish between material that is above cut-off grade and material that is below cut-off grade by estimating recoverable reserves. Because drilling does not cover the entire area to be mined, recoverable resource estimation attempts to predict the quality (grade) and quantity (tonnage) from a limited number of data points (Gulule, E. P., 2016).

According to an estimation of economic losses due to poor blast hole sampling in open pits, errors in sampling and preparation as well as estimation methodology are to blame for losses of the order of millions of dollars annually. The estimation methodology is responsible for larger losses than those related to sampling errors. The use of the polygonal method instead of kriging generates invisible profit losses. Kriging is less sensitive to the level of sampling error than the polygonal method. Other techniques, such as geostatistical simulations, could be evaluated to improve the SMU classification. The total profits can be significantly impacted by improvements like the adoption of geostatistical estimation methods, better equipment for sampling and sample preparation, and staff training (Magri & Ortiz, 2000). Geostatistical simulation allows for the quantification of losses generated by poor blast hole sampling and imperfect estimation.

Inverse Distance Weighting methods assume that samples taken close together will have more characteristics in common than samples taken farther apart, with anisotropies not frequently considered when using the same corner point grade and thickness for multiple orebodies. Kriging is superior to the IDW method because it considers not only distance but also spatial variability, as well as sample redundancy and proximity. Outlier values in geostatistics can result in distorted variograms due to the high nugget effect. The main aim of cutting the high-grade values is to alter the samples' distribution.

Kriging uses a set of simultaneous linear equations for each point on the output grid such that all the actual input data is optimally weighted according to distance using the semi variogram. Kriging is the geostatistical estimation method developed to provide the optimal linear and unbiased estimates. It depends on expressing spatial variation of the property in terms of the variogram (or correlogram), and it minimizes the prediction errors, which are then estimated. The technique is based on the assumption that the variable to be estimated is a regionalized variable. The technique is used if the underlying conditions of second-order stationarity are met, which means that the sample data's mean and variance stay unchanged in space at a minimum. Ordinary kriging is the most widely used kriging method. It serves to estimate a value at a point of a region for which a variogram is known, using data in the neighborhood of the estimation location. Ordinary kriging can also be used to estimate a block value (Wackernagel H., 1995).

3. Case study

3.1. Background

The case study's data came from a copper mine in Chile, South America. This dataset contains 28,634 blasthole data points collected from a 15 m bench height open pit. The copper grade reported has a mean, standard deviation, and coefficient of variation of 0.841, 0.330, and 0.392, as shown in the histogram in Figure 1.

Sorting the data by elevation resulted in the selection of a bench. The bench data chosen ranged between 270 and 285 meters. On the chosen bench, 1855 blasthole data were discovered. According to the histogram in Figure 2, the mean, standard deviation, and coefficient of variation of the copper grade reported are 0.815, 0.203, and 0.249, respectively. Figure 3 shows the location of the blast hole as well.

The following GSLIB functions will be used to simulate blasthole data: **nscore**, **vmodel**, **sgsim**, and **blkavg**. **Histplt** and **pixelplt**, on the other hand, will be used to visualize the output files.

The normal score values for the numerous blast holes are calculated. The normal score feature alters the dataset to closely resemble a standard normal distribution. This is performed by comparing the ranks provided by a normal distribution to the ranks acquired by ordering the values in the dataset from lowest to highest. When the word "normal score" is used, it is typically expected that the result can be compared to a table of standard normal probability. Figure 4 depicts the normal score values. A typical 8-directional variogram with 40 lags, 10m lag separation distance, and 5m lag tolerance is generated from the normal score output file, as shown in Figure 5.

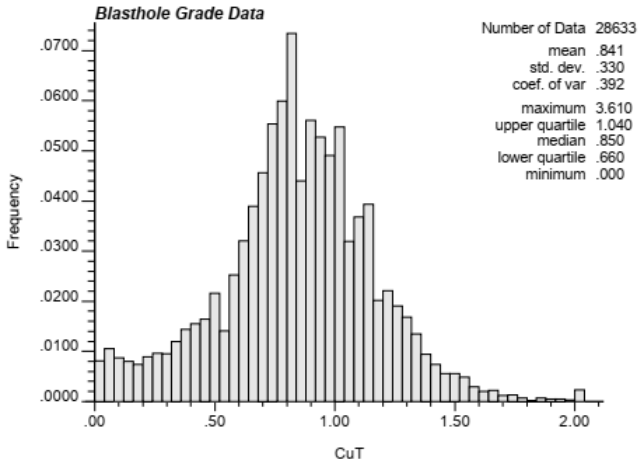


Figure 1: Histogram of the grade of all the blast hole data obtained from the mine.

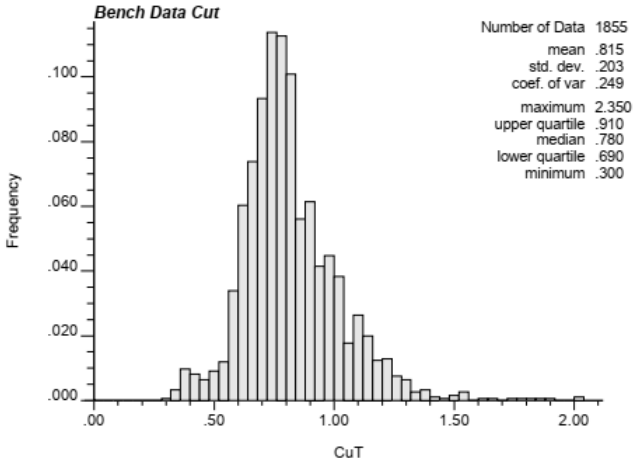


Figure 2: histogram of the grade of the selected blast hole data.

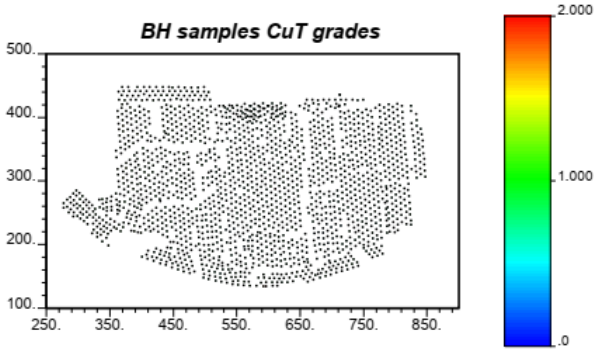


Figure 3: map of blast hole samples for Cu grade.

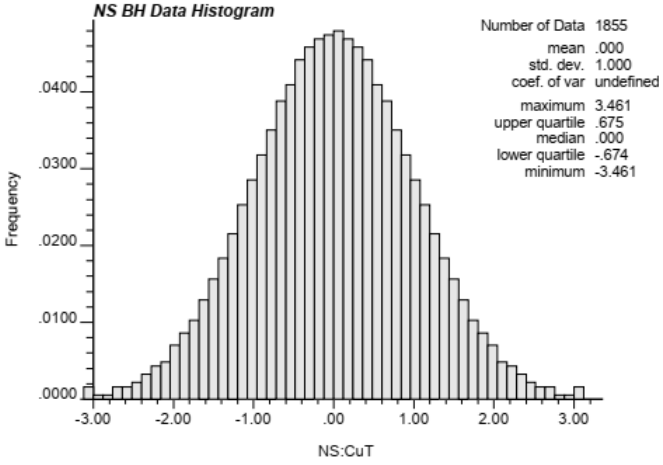


Figure 4: Histogram showing the normal score of the selected data.

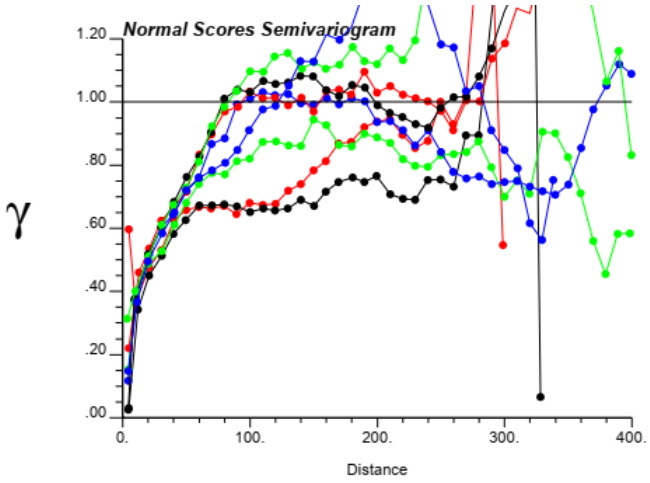


Figure 5: normal score directional variograms.

The second and sixth azimuth directions of 22.5 and 112.5 fall in the outer section of the normal score semi variogram shown in Figure 5. A variogram model is then created for the two dimensional variogram. Figure 6 depicts this.

The selected blasthole data is then utilized to produce a simulated model using sequential Gaussian simulation with the corresponding variogram model. SGSIM, a stochastic method, was developed to avoid the smoothing effect caused by deterministic methods by generating a number of stochastic realizations. Actual data from sampled sites and values from previously simulated locations are used in the sequential procedure to inform each unknown location. Furthermore, because distinct random paths are constructed that can pass through the unsampled sites in different orders, SGSIM can generate a number of equally plausible outcomes to investigate and evaluate the uncertainty (Verly, 1993). The simulation was implemented with a maximum search radius of 50m in all directions. A single simulation with a minimum and maximum original data of 8 and 12, respectively, is used as a representation of the ground truth distribution of grades at point support. Figure 7 depicts this simulated scenario.

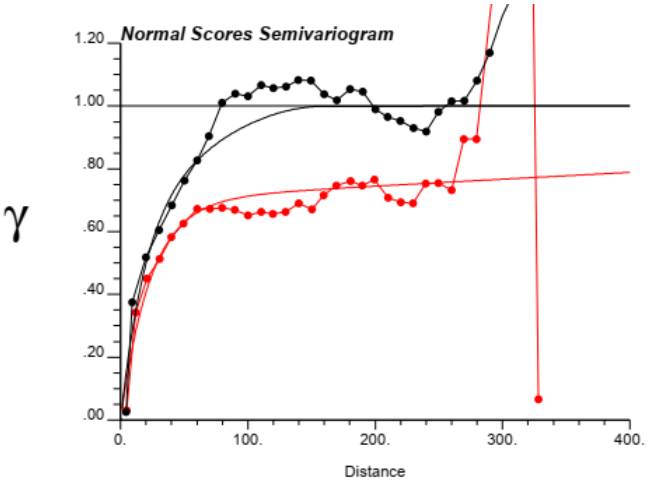


Figure 6: Normal Score variogram model.

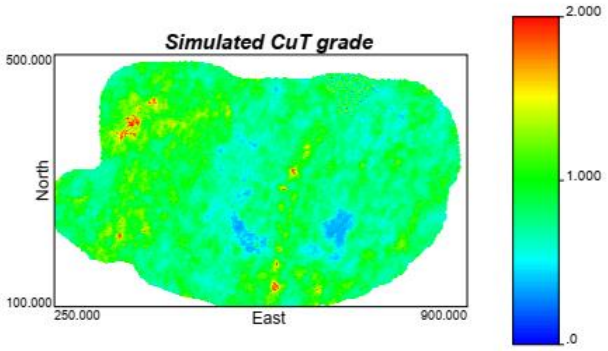


Figure 7: Simulated point Cu grades.

The ground truth at point support is used to generate the block ground truth, by averaging the simulated points inside each 10 x 10m block. The Ground Truth Cut block grades are shown in Figure 8.

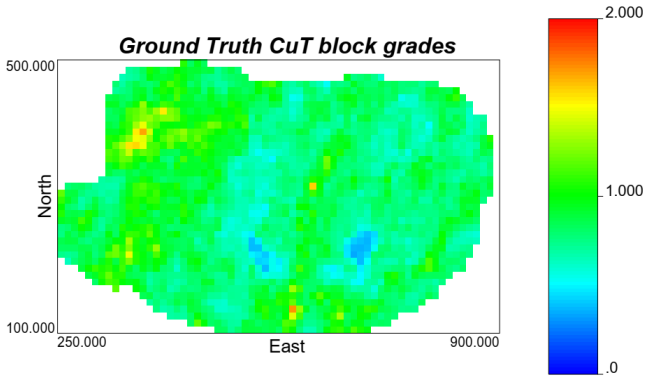


Figure 8: Simulated block Cu grades.

3.2. Addition of errors and block estimation

Ordinary Kriging is preferred in block estimation because, despite the assumption that the mean is unknown, it assumes stationarity on the neighborhood of the estimate point. Estimation is performed over 10 x 10m blocks and using a 4 x 4 block discretization, with minimum and maximum data for kriging of 4 and 16, respectively. Ordinary kriging is conducted using the previously developed variogram model. Figure 9 depicts the estimated block after using Ordinary Kriging.

Errors are introduced in the copper grade recorded in the blasthole data. These errors are used to assess the effect of precision of the data as well as the consequences of having biased data. A random error is added to the blasthole data to evaluate precision when 10%, 20%, 30%, and 50% errors are introduced. For the bias scenarios, copper grades of 110%, 90%, 130%, 70%, 150%, and 50% of the true blasthole sample value are calculated.

Ordinary kriging is then used to estimate the block grades. Negative grades that appear during the estimation are made zero. The estimated block models resulting from using samples with precision errors

are shown in Figure 10, while those resulting from using samples with bias are seen in Figure 11, showing the different scenarios. A statistical summary of the various errors is shown in Table 1 and Table 2.

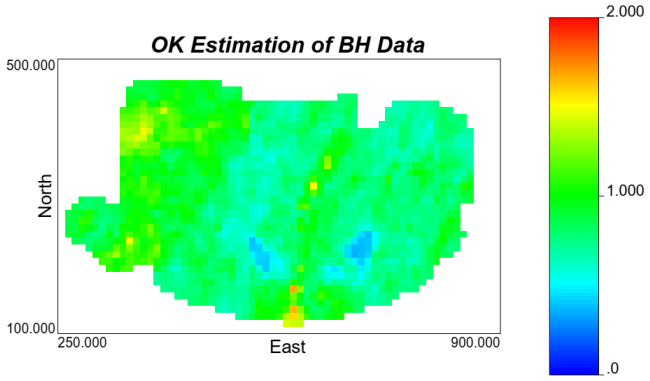


Figure 9: Ordinary kriging estimation of Cu grades at block support.

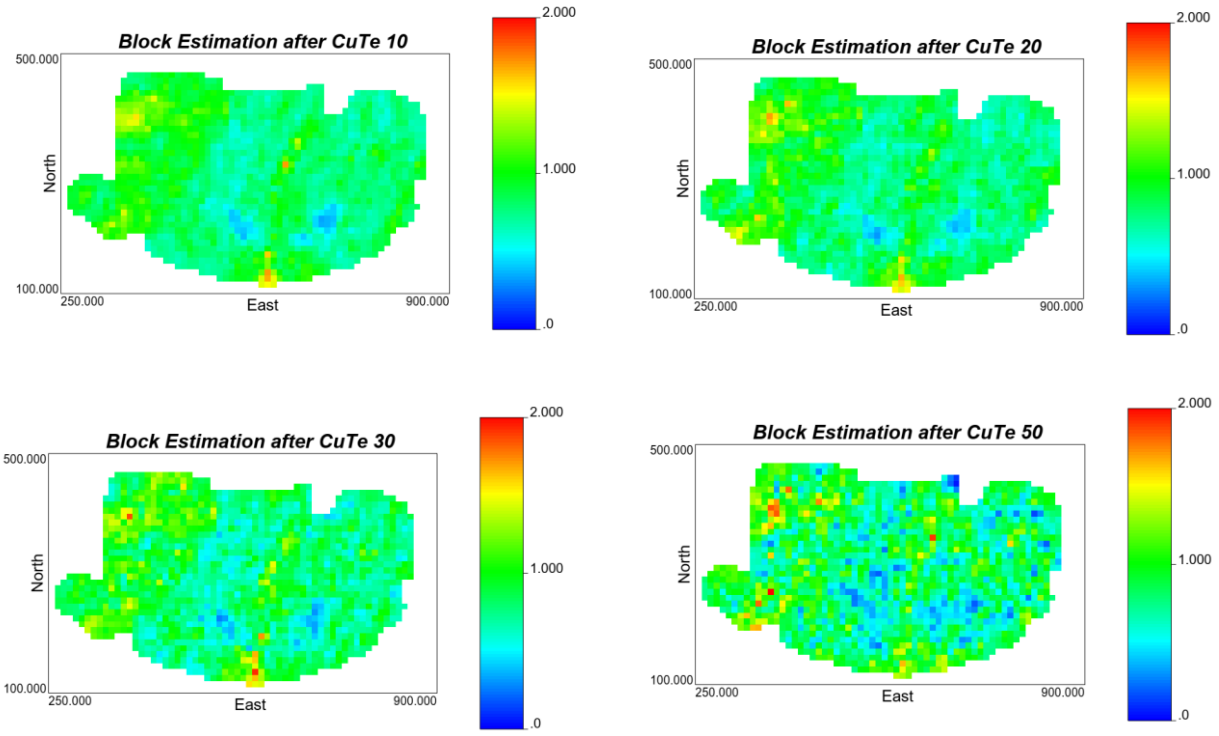


Figure 10: Block estimates based on samples with added precision errors of 10% (top left), 20% (top right), 30% (bottom left) and 50% (bottom right).

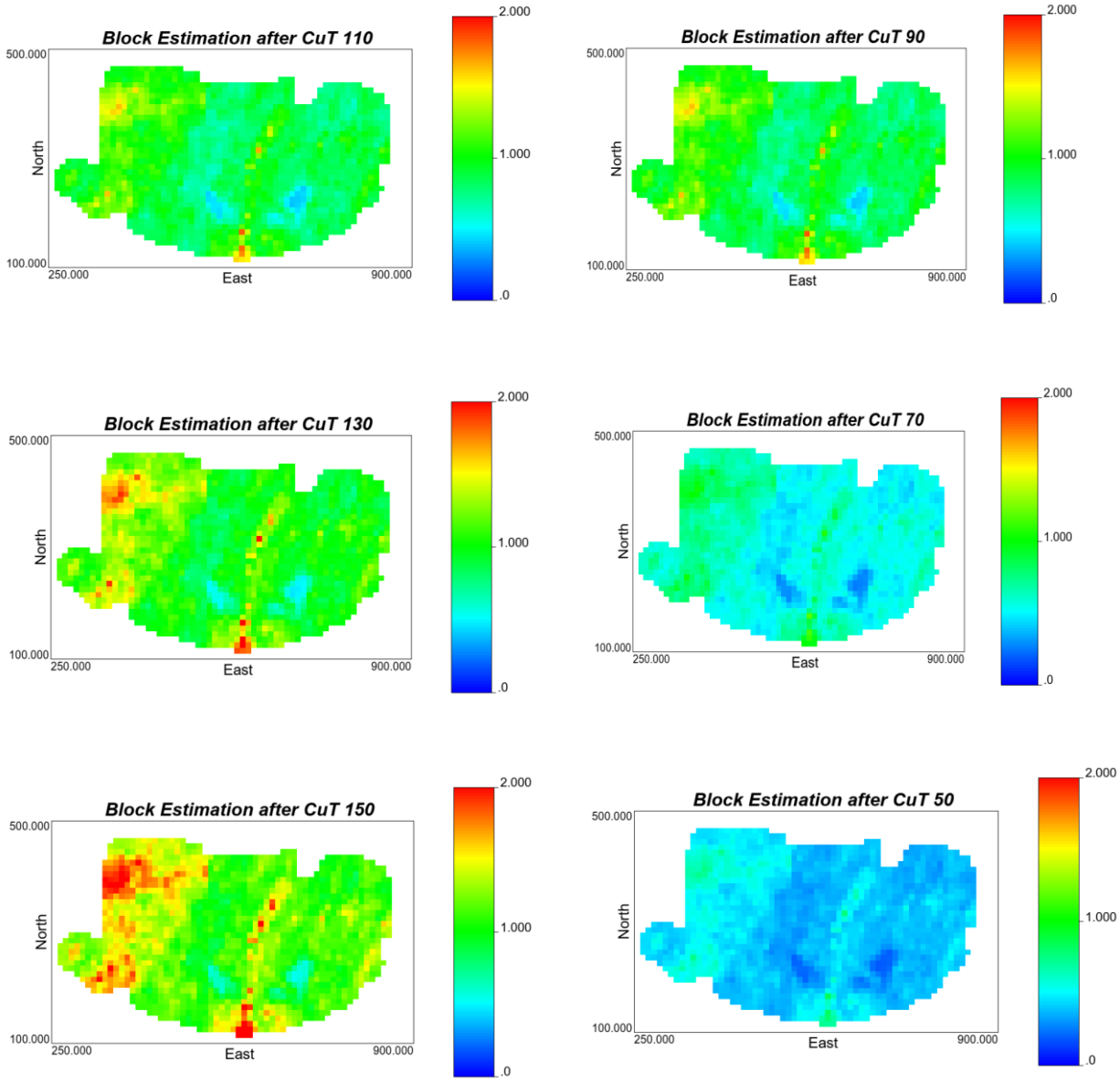


Figure 11: Block estimates based on samples with added bias of +10%, +30% and +50% (left column), and -10%, -30% and -50% (right column).

Table 1: statistical summary of ground truth point and block simulated values, and original blasthole samples and samples with added error for different precisions.

Parameters	SIM point	SIM block	BH data	Prec. 10%	Prec. 20%	Prec. 30%	Prec. 50%
number of data	49005	2037	1855	1855	1855	1855	1855
mean	0.820	0.820	0.815	0.817	0.820	0.811	0.795
standard deviation	0.188	0.165	0.203	0.220	0.265	0.323	0.454
coefficient of variation	0.229	0.202	0.249	0.269	0.323	0.399	0.560
maximum	2.350	1.711	2.350	2.207	2.282	3.260	2.851
upper quartile	0.910	0.904	0.910	0.924	0.962	0.988	1.066
median	0.790	0.797	0.780	0.782	0.787	0.775	0.760
lower quartile	0.700	0.714	0.690	0.675	0.640	0.594	0.477
minimum	0.181	0.329	0.300	0.281	0.219	0.000	0.000

Table 2: statistical summary of blasthole samples with bias.

Parameters	SIM point	SIM block	Bias +10%	Bias -10%	Bias +30%	Bias -30%	Bias +50%	Bias -50%
number of data	49005	2037	1855	1855	1855	1855	1855	1855
mean	0.820	0.820	0.897	0.734	1.060	0.571	1.223	0.408
standard deviation	0.188	0.165	0.223	0.182	0.263	0.142	0.304	0.101
coefficient of variation	0.229	0.202	0.249	0.249	0.249	0.249	0.249	0.249
maximum	2.350	1.711	2.585	2.115	3.055	1.645	3.525	1.175
upper quartile	0.910	0.904	1.001	0.819	1.183	0.637	1.365	0.455
median	0.790	0.797	0.858	0.702	1.014	0.546	1.170	0.390
lower quartile	0.700	0.714	0.759	0.621	0.897	0.483	1.035	0.345
minimum	0.181	0.329	0.330	0.270	0.390	0.210	0.450	0.150

3.3. Cutoff grade application

In analyzing the effect of the errors introduced, a cutoff is applied. To avoid issues related to mine and processing capacity, the cutoff selected is close to the breakeven cutoff grade, in this case 0.554% Cu. The ore blocks estimated grades are used to decide to send the block to the processing plant, while the actual grade (the one from the block support simulated ground truth) is used to assess its value. Since block grade estimates are based on samples and these can have sampling errors related to precision or bias, the ore/waste classification will be imperfect, which has an impact on the overall economic value.

In calculating the tonnage of the blocks a bulk density of 2.7 t/m^3 is assumed. The height of block is 15 m with the block area of 10m x 10m. The metal content for all the scenarios is derived from the multiplication of the grade and tonnage of the block and a recovery of 90% is assumed.

3.4. Economic evaluation of the blocks

This computation determines the profit or loss of each block and bench under discussion, taking into account the costs to be incurred and compensations to be made for the cutoff grade used.

The blocks with an average grade higher than the cutoff grade are termed ore, which is computed by deducting the cost of processing from the value of the block in terms of its grade. Waste blocks give no economic advantage. These results are added together to determine how much profit or loss will be earned in the sequence of scenarios mentioned to check for precisions and biases. Notice that both ore and waste blocks need to be removed from the bench, so the mining cost is not considered in these calculations.

The values of all parameters used are based on assumptions about a real-world scenario. The parameters used in the evaluation are shown in Table 3. The grade assigned to each estimated and error-added block is based on the simulated grade, since this constitutes the true grade of the block and determines the profit or loss actually incurred by the operation. The decision about where to send the block, on the other hand, is made on the estimated grade, which may have been obtained from samples with error. If the estimated grade of the block is greater than the cutoff grade, the block becomes an ore block, with its economic evaluation based on the simulated grade of the block. However, if the grade of block estimated is less than the cutoff grade, the block becomes a waste block, for which the economic evaluation is the cost of mining the entire block. Each block was estimated to be 10m x 10m x 15m in size. A bulk density of 2.7 was used in determining the tonnage of the block, which is denoted by the block volume multiplied by the bulk density. The outcome of the evaluation is shown in Table 4 and Table 5 for the precision errors and biases. A graphical representation of the outcome of the evaluation is shown in Figure 12 and Figure 13.

Table 3: economic parameters used in the profit calculations.

Parameters	Value	Units
Price	3.3	USD/lb
Recovery	90	%
Processing cost	22	USD/t
Ton	10x10x15	Ton/block
Metallurgical cost	1.3	USD/lb

Table 4: ore blocks, their grade, tonnage, metal and profit, along with their relative error with respect to the ground truth, for different scenarios of blasthole samples precision errors.

Cutoff 0.554%Cu Parameter	SIM block	BH	Prec. 10%	Prec. 20%	Prec. 30%	Prec. 50%
Ore Blocks	1526	1538	1531	1520	1448	1297
Average grade	0.83	0.83	0.83	0.83	0.84	0.84
Tonnage (kT)	6180	6229	6201	6156	5864	5253
Metal (M lb Cu)	113	114	114	113	108	98
Profit (M$)	68.0	68.0	67.9	67.6	66.0	60.1
Grade error (%)	**0.0**	**-0.3**	**-0.2**	**-0.1**	**0.8**	**1.3**
Tonnage error (%)	**0.0**	**0.8**	**0.3**	**-0.4**	**-5.1**	**-15.0**
Metal error (%)	**0.0**	**0.5**	**0.2**	**-0.5**	**-4.4**	**-13.9**
Profit error (%)	**0.0**	**-0.1**	**-0.2**	**-0.6**	**-3.0**	**-11.6**

Table 5: ore blocks, their grade, tonnage, metal and profit, along with their relative error with respect to the ground truth, for different scenarios of blasthole samples bias errors.

Cutoff 0.554%Cu Parameter	SIM block	BH	Bias +10%	Bias -10%	Bias +30%	Bias -30%	Bias +50%	Bias -50%
Ore Blocks	1526	1538	1548	1488	1562	811	1576	1526
Average grade	0.83	0.83	0.83	0.84	0.82	0.94	0.82	1.20
Tonnage (kT)	6180	6229	6269	6026	6326	3285	6383	421
Metal (M lb Cu)	113	114	114	111	115	68	115	11
Profit (M$)	68.0	68.0	67.9	67.8	67.7	50.1	67.4	10.8
Grade error (%)	**0.0**	**-0.3**	**-0.5**	**0.7**	**-0.9**	**12.9**	**-1.4**	**44.2**
Tonnage error (%)	**0.0**	**0.8**	**1.4**	**-2.5**	**2.4**	**-46.9**	**3.3**	**-93.2**
Metal error (%)	**0.0**	**0.5**	**0.9**	**-1.8**	**1.4**	**-40.0**	**1.8**	**-90.2**
Profit error (%)	**0.0**	**-0.1**	**-0.2**	**-0.4**	**-0.5**	**-26.3**	**-1.0**	**-84.2**

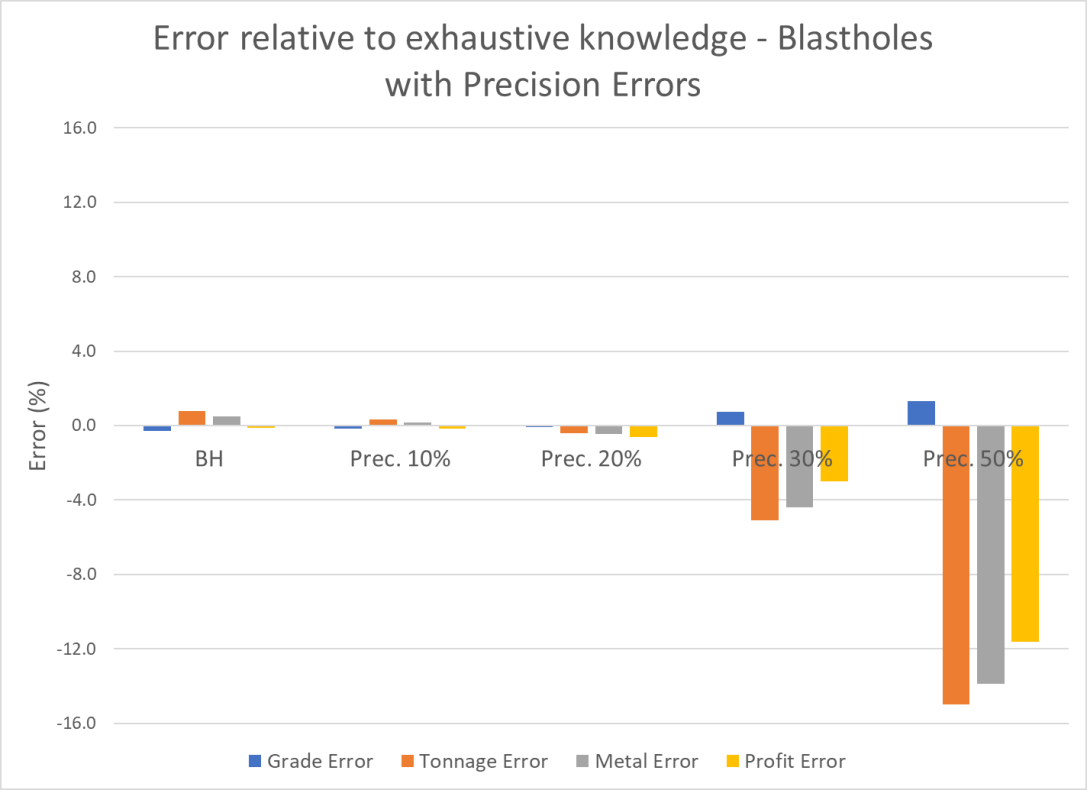


Figure 12: relative errors in grade, tonnage, metal and profit for the case of blasthole samples with precision errors.

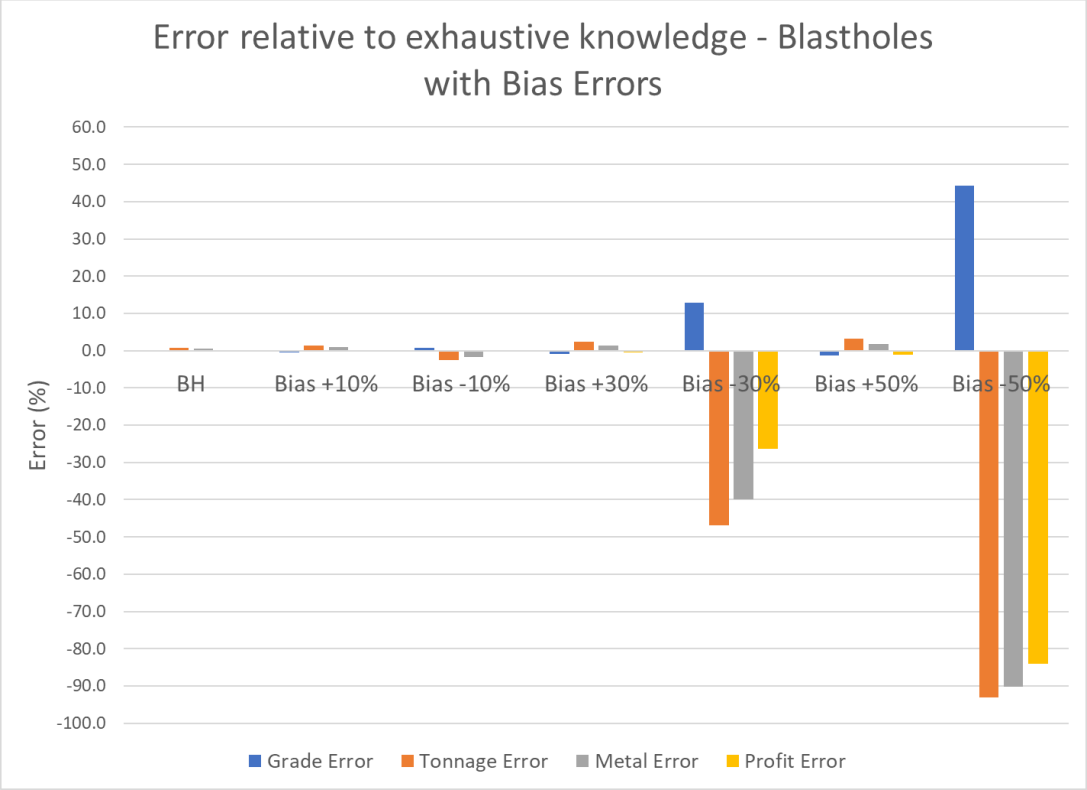


Figure 13: relative errors in grade, tonnage, metal and profit for the case of blasthole samples with bias errors.

3.5. Results discussion

The results demonstrate that the use of samples with errors lead to a short term block model that misclassifies the correct destination of ore and waste blocks. This translates into a loss of profit.

The use of unbiased and precise samples at blastholes, along with an unbiased and optimum estimation technique such as ordinary kriging, generates a model that is very close to the perfect classification of blocks, which in practice is unachievable, since we do not have access to the true block grades. In this work, that "reality" is based on one simulated model, developed at point support and then block averaged to represent the true grade of blocks. The points simulated are conditioned to the actual blasthole dataset of the mine. The classification using the blasthole samples without any added error, generates relative errors in the ore grade, tonnage, metal content and profit lower than 1%. Profit is only 0.1% lower than the unachievable case of perfect knowledge. This is encouraging and should suggest to practitioners that ordinary kriging must be used in short term planning and during grade control to determine the block grades prior to their assignment as ore or waste.

When samples suffer from precision errors, their values are overall unbiased, but at any location they can be slightly higher or lower than the actual value. Depending on the level of noise, the estimation of the block grades will be affected proportionally, and this means more blocks are incorrectly estimated above or below the cutoff. Blasthole sampling is known to be of poor quality in many mining operations. However, quality control and quality assurance procedures ensure that it will rarely exceed 15 to 20% of relative error (this is the error relative to the mean value of the sample grades). It can be seen that when

the error in precision is 10 or 20% in the blasthole samples, the estimated blocks will be misclassified more often. However, the relative errors in grade, tonnage, metal content and profit are still below 1% with respect to the case of perfect knowledge. Profits decrease in 0.2 and 0.6% for 10% and 20% precision errors. However, when the samples precision is lower, with errors of 30 and 50%, the impacts on profit explode quickly. Profit falls by 3.0 and 11.6% for these two cases, which represents a loss of $2 million and $8 million, respectively. This is a loss over just over 6 million tons of ore. In a large open pit, this could represent a month of production, that is, $2 million or $8 million per month.

When samples are biased, losses can be very significant, depending on the cutoff and the direction of the bias. A positive bias will send more material to the mill, which in the end may not generate too much of a loss, if the "waste" is not too low in its true grade. On the other hand, if the bias is negative, many blocks of ore will be sent to the waste dump, since their grade was predicted with a systematically lower value. The few blocks that are correctly assigned to the processing plant will not generate enough metal to compensate for the lost ore. We see significant profit losses for biases of -30% and -50%. Losses in other cases are minor, although they are always there.

4. Conclusions

The study shows that ordinary kriging is a great tool for forecasting the true block grades and the profit losses due to the lack of perfect knowledge of the true grades is very small. When samples are affected by sampling errors, the effects can be tremendous in profit. Precision issues may lead to million dollars losses over a year's production. Furthermore, when samples are biased, this may lead to a loss of ore, which will be misclassified and sent to the waste dump. Comparison with other estimation methods would provide insight with respect to their unbiasedness, optimality and robustness. Also, accounting for mining and processing capacity constraints would change the results. These possibilities could be explored in future work.

5. References

Abzalov, M. Z., Menzel, B., Wlasenko, M., & Phillips, J. (2010). Optimisation of the grade control procedures at the Yandi iron-ore mine, Western Australia: Geostatistical approach. *Applied Earth Science*, *119*(3), 132–142. https://doi.org/10.1179/1743275811Y.0000000007

Deutsch, C. V., Magri, E., & Norrena, K. (2000). Optimal Grade Control Using Geostatistics and Economics. *Society for Mining, Metallurgy, and Exploraction Inc.*, *308*.

Glacken, I. M. 1997. Change of support and use of economic parameters for block selection, in Geostatistics Wollongong '96, (eds. E. Y. Baafi and N. A. Schofield), Vol. 2, 811–821, Dordrecht, Netherlands, Kluwer

Gulule, E. (2016). Comparative Analysis of Ordinary Kriging and Sequential Gaussian Simulation for Recoverable Reserve Estimation at Kayelekera Mine. https://core.ac.uk/download/pdf/188769486.pdf

Magri, E. V., & Ortiz, J. (2000). *Estimation of Economic Losses Due to Poor Blast Hole Sampling in Open Pits*. 11.

Verly, G. (1993). Sequential Gaussian Simulation: A Monte Carlo Method for Generating Models of Porosity and Permeability. In A. M. Spencer (Ed.), *Generation, Accumulation and Production of*

Europe's Hydrocarbons III (pp. 345–356). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-77859-9_28

Verly, G. (2005). Grade Control Classification of Ore and Waste: A Critical Review of Estimation and Simulation Based Procedures. *Mathematical Geology*, *37*(5), 451–475. https://doi.org/10.1007/s11004-005-6660-9

Wackernagel, H. (1995). Ordinary Kriging, access on November 10, 2022, from https://link.springer.com/chapter/10.1007/978-3-662-03098-1_11

Review of blast movement measurements for grade control[1]

Noble E. Potakey (n.potakey@queensu.ca)
Julian M. Ortiz (julian.ortiz@queensu.ca)

Abstract

Blasting is carried out in mining operations to break down rocks and to maximize material movement. In open pit mines, this invariably involves huge amount of explosive energy which causes rock materials to be displaced from their original position. This movement is detrimental to the accurate delineation of the predefined ore and waste zones and could lead to ore loss and dilution if not accounted for. Direct measurements such as the use of visual markers have been widely patronized in most campaigns. Sandbags retrieved after blasting show that pre-blast grades could be displaced up to 10-15 meters after blasting. Blast movement monitors (BMM) developed by a group of researchers from the University of Queensland currently provide the most accurate method of blast-induced rock movement despite the cost of data acquisition. In recent years, indirect determination of blast movement has been advocated using software and complicated simulation algorithms. In this paper, the limitations of direct blast movement techniques as well as the feasibility of indirect measurement models are discussed. Considering that there is no easy-way and cheap method to determine post-blast ore boundary, a machine-learning (ML) approach and a corresponding evaluation system have also been proposed in the literature.

1. Introduction

After tremendous work has been done to define and model the distribution of minerals in a rock mass, the rock undergoes a comminution process before the mineral is extracted. For most scenarios, the first stage of the comminution process is blasting, and this allows efficient excavation and haulage after the rock has been fragmented. Blasting is done using explosives inserted into holes drilled in the rock. Upon detonation, the chemical energy in the explosive is released, and the solid explosive becomes transformed into a pressurized gleaming gas that shatters and move rocks in the path of least resistance resulting in a muck pile (Hustrulid, 2011). Lawrence (1944), Thornton (2009) and Zhang (2016) provide more information on the detonation theory and the mechanics of rocks breakage.

Considerable amount of research has been done on blast optimization, but this has often been in the area of ground vibration, rock damage, fragmentation, blast design, strain energy, and in regard to environmental safety (Blair & Minchinton, 1997; Persson, 1997; Sołtys et al., 2017; Zou, 2017). The impact of blast-induced rock movement on predetermined grade distribution of the rock has not been extensively explored and leaves much to be desired (Thornton, 2009). Grade control is a compendium of procedures and practices aimed at sending the mined material to the right destination and that involves considering the post blast movement of rocks. Disregarding this situation will lead to misclassification, that is,

[1] Cite as: Potakey N. E., Ortiz J. M. (2022) Review of blast movement measurements for grade control, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-06, 85-95.

mistaking ore for waste or waste for ore; a mixing of low grade and high grade materials; sulfides to oxides identification issues; or other contaminants – collectively referred to as ore loss and dilution (Rosa & Thornton, 2011). Figure 1 is an illustration of how ore loss and dilution occur due to ore block movement.

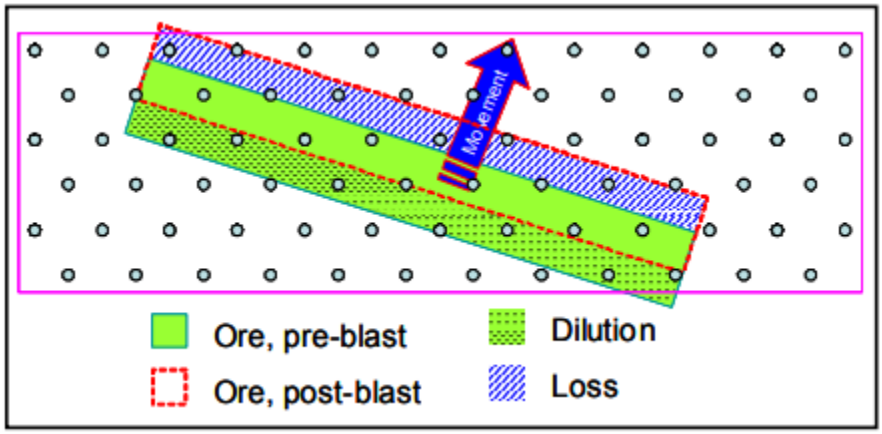


Figure 1 Ore loss and dilution during blast. Source: (Thornton et al., 2005)

To account for this movement, various methods have been used to measure or model the pre and post blast rock locations. Two main approaches used to measure or model blast movement are the direct measurements based on the use of physical markers, and indirect measurements such as numerical modelling. Direct measurement method involves the use of objects inserted into the pre blast rock and their post blast location are retrieved after excavation and measured (Rosa & Thornton, 2011). The use of simple visual markers such as sandbags and poly-pipes have been employed in several open pit mines to track rock movement because it is simple and inexpensive. However, its limitations include a low turnout of recovered markers and its inability to provide a three dimensional movement pattern (Thornton, 2009). Another direct method used in blast movement measurement is the application of remote sensing devices (Vasylchuk, 2019). Developed by a group of researchers from the University of Queensland, Australia, the electronic blast movement monitor (BMM) quickly became a grade control to measure rock movement. The BMM device relies on transmitters that are installed in the blast prior to blasting, which are recovered after the blast by a special detector and the data is processed with a software.

Modelling the entire blasting process as an alternative for direct measurements have received a mixture of feedback even though it is a good prospect to monitoring blast movement. Lack of complete knowledge about the geological domain, location of rock breakage and mechanical properties of the rock, together with uncertainty in blast parameters undermines the accuracy of the model (Vasylchuk & Deutsch, 2019). Considering that the BMM method is expensive, and most companies cannot afford it, calibrating a numerical model that could measure rock movement is worthy of research (Vasylchuk & Deutsch, 2018).

In the next sections of this paper, we review the blast movement methods in operation in most mines according to case study, highlighting their limitations and the feasibility of indirect measurement approaches. Subsequently, a novel machine learning approach is discussed as an indirect method considering that there is no easy-way and cheap method to determine post-blast ore boundary.

2. Blast Movement Measurement

Understanding material movement during a blast has always been an intriguing area to mine operators especially where there is no clear visual distinction between ore and waste. Various methods used have demonstrated a mixture of success and some limitations. Traditional methods of understanding the movement of the rock was to compare and contrast pre and post blast topographic surfaces (Vasylchuk & Deutsch, 2018). In context, blast movement measurement has been categorized into i) direct measurements and ii) indirect measurements.

2.1 Direct Measurement of Blast-induced Rock Movement

This type of measurement involves the use of physical markers to track material movement. Two major direct approaches have been used: i) the use of visual markers, and ii) using remote sensing devices.

2.1.1 Visual Markers

The use of visual markers encompasses objects such as sandbags, chains or pipes inserted into the rock before blasting and their post blast location identified and measured (Rosa & Thornton, 2011). In their research, Taylor (1995) and Zhang (1994) appraised the use of sandbags and wooden stakes as markers for rock displacement during blasting. Results indicate that even though these visual markers are simple, cheaper, and relatively accurate, only about forty percent (40%) of the markers were recovered and it took several days for all the bags to be found. A more common industrial approach is the use of plastic pipes inserted into additional holes drilled within the blast area. And as the pipes are exposed during excavation, their locations are surveyed. For bench-by-bench excavation, the process is repeated for each level. Figure 2 shows an example of a recovered pipe after blasting.

However, the disadvantage of this method is that the data for processing is not available until the markers are found and the ore has been excavated (Fitzgerald et al., 2011). This does not allow the proper design and adjustment of dig polygons prior to excavation. The use of the poly pipes also presents several limitations including the generation of only two-dimensional vector measurements, poor recovery of pipes for lower-level benches and it being labor-intensive. It must however be noted that the use of markers is an ad-hoc approach, and not many of such are published in literature.

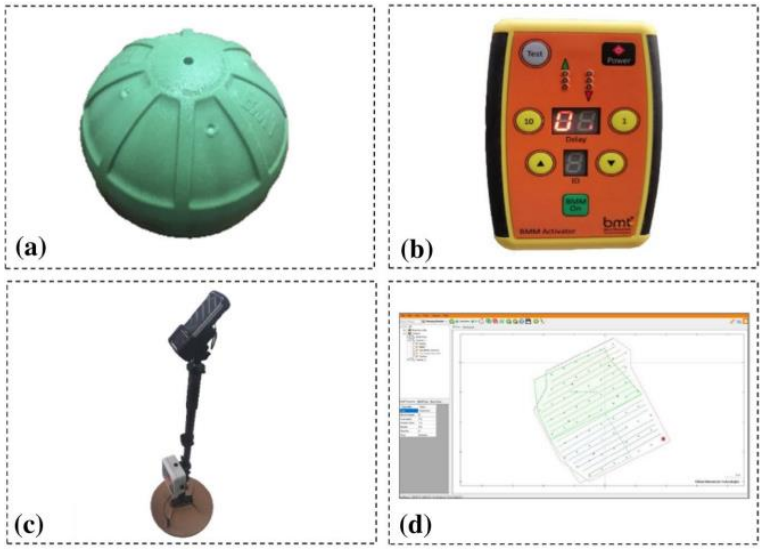
Polypipe exposed after excavation.

Figure 2 Pipe recovered after blasting. Source: (Rosa & Thornton, 2011)

2.1.2 Remote Sensing Devices

A modern approach to directly measure blast movement is the remote detecting equipment. This is an electronic method that aims at alleviating some of the limitations of the visual methods such as reducing the arduity. In remote sensing methods, metallic or magnetic targets are used instead of marker bags or pipes, and their post blast locations identified using remote sensing or electronic devices. Various methods have been tested including *Ground Penetrating Radar*, *Magnetometry*, *Metal detection and recently, the Radio frequency (RFID) tags* (Thornton, 2009). However, most suffer limitations such as damage of targets by excavators, use of only one target in each hole and targets must be placed close to the surface or on the surface, which is detrimental to accurately measure movement dynamics.

By far, a remote sensing approach that has proven very effective and is almost the most accurate method of blast-induced rock movement monitoring is the blast movement monitoring (BMM) device. This method is used in mines such as the Husab Uranium mining project in Namibia, the second largest world producer of uranium (Yu, Shi, Zhou, Rao, et al., 2019). Developed by a team of researchers from the University of Queensland and later commercialized under the Blast Movement Technologies (BMT), the BMM system comprises of transmitters that are installed in separate holes drilled between blastholes and are held in place by drill cuttings or stemming (Fitzgerald et al., 2011). After the blast, the transmitters are located with a special detector and the data is processed with a purpose-designed software. The structure of a modern BMM device is shown in figure 3.



(a) ***BMM ball:*** *contains transmitters that emit electromagnetic signals which is detected by the BMM detector.*

(b) ***BMM activator:*** *activates the BMM ball prior to the blast.*

(c) ***BMM detector:*** *locates the BMM ball after the blast.*

(d) ***BMM explorer:*** *computer software that does the post blast ore boundary calculations.*

Figure 3 The blast movement monitoring system. Source: (Yu, Shi, Zhou, Rao, et al., 2019).

The BMM ball can be detected to a depth of around 25 m after blasting. Once the horizontal location of the ball is pinpointed, the signal is recorded to determine the depth below the surface, and then the three-dimensional (3D) movement vectors is calculated. The 3D movement vectors obtained is then applied to the ore block boundaries determination by the system software with results usually ready within an hour or two after the blast (Adam & Thornton, 2004). With a battery life of 12 hours and more, and excellent detection rates of about ninety per cent (90%), the BMM system has proven to be very effective and practical for grade control.

2.2 Indirect Measurement of Blast-induced Rock Movement

Indirect blast movement measurement methods became necessary to complement the direct measurement method in saving time and money. Visual markers are labor intensive and BMMs are not cheap. So an indirect measurement is suggested which involves the use of algorithms and software to infer the movement of rocks based on data and other field parameters collected (Vasylchuk & Deutsch, 2018). In most cases, the post blast topographic surface is traced and compared to the pre-blast topography and an approximate blast movement model is developed. In this section, we will look at numerical simulated models and machine learning (ML) models.

2.2.1 Numerical modeling of blast movement

Early attempts to numerically model blast movement was hindered by computational capability. Early blast movement models developed included the Universal Distinct Element Code (UDEC) by Cundall (1980) which attempts to model by simulating behavior of jointed rock masses subjected to high and transient loadings; the Block and Bump model by Schamaun (1986) where blast movement is represented by blocks and circles and where the dynamic movement of rock particles is controlled by parameters such as the geological characteristics of mine benches, shapes and sizes of the particles, and cohesive forces between rock particles. The advanced Distinct Motion Code (DMC) model presented by Preece et al. (1997) allowed the incorporation of the properties of explosives for modeling the motion of rocks.

Having mentioned that, in recent years, simple and efficient models have been developed such as the simple blast movement model by Furtney et al. This model illustrates among other things, how the chemical energy of the explosive is distributed during blasting and how it impacts the displacement of rocks. The model seems able to predict the face velocities using generic rock properties as inputs within a certain degree of accuracy. Detail of this work can be found in Furtney et al. (2013). In 2018, Vasylchuk and Deutsch described a blast movement model using pre and post blast topographic features. In the model, the algorithm proposed translated the the pre-blast grid locations to post-blast locations, and a 3D model of the post blast muck pile was created. Post-blast locations were inferred from discretized pre-blast locations. Figure 4 shows the pre and post blast models generated by their algorithm.

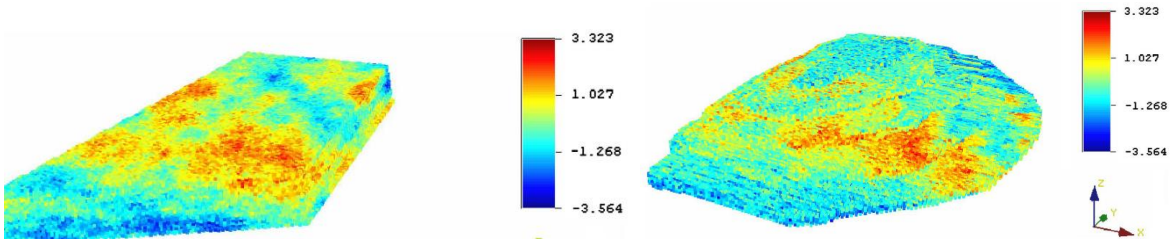


Figure 4 Pre (left) and post (right) blast 3D models with assigned grades Source: (Vasylchuk & Deutsch, 2018)

Vasylchuk and Deutsch (2019) advanced their research by developing an empirical optimization algorithm that incorporates the use of direct measurements to topographic monitoring. Results from a fabricated scenario demonstrated the model's ability to map pre-blast grade onto post-blast muck pile within a reasonable time and still honored real information about blast movement. Figure 5 shows the grade distributions prior to and after blasting by the model and their ultimate destinations.

Despite the interesting approach of numerically modelling blast movement, it also draws legitimate concerns. Some of which are the uncertainties in the blast parameters, lack of absolute knowledge of the geological features, fracture locations and mechanical properties of the rock (Vasylchuk & Deutsch, 2019). According to Yu et al. (2019), theoretical calculations and numerical simulations do not provide accurate blast-induced rock movement measurements. The discrepancy between a modelled and a measured blast movement was tested by Rosa and Thornton (2011) and the error margin was from 1 to 7 meters which is estimated to be equivalent to a loss of about 2.2 to 4.8 million dollars. They further suggested that blast models should be validated with actual pre and post blast bench configurations.

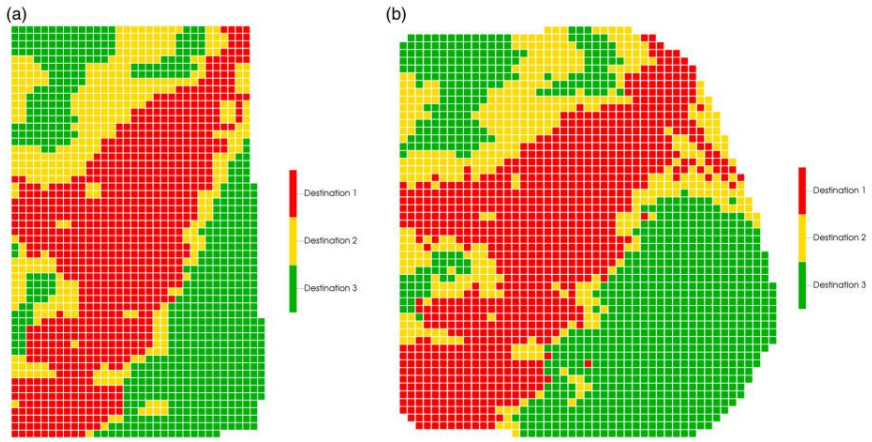


Figure 5 Pre (a) and post (b) blast classification of materials Source: (Vasylchuk & Deutsch, 2019).

2.2.2 Machine Learning (ML) approach

The advancement of computers and technology has aided the processing and manipulation of high volumes of data within the shortest possible time. Machine learning algorithms (MLA) are a collection of advanced statistical tools to provide a faster and better way of processing data using high-level processors. The application MLA has received a lot of successes such as the application of Artificial Neural Networks for grade estimation in mineral resource estimation (Abuntori et al., 2021). New methods such as deep networks performed excellently in its predictive ability with both structured and unstructured data (Shen et al., 2018). Random Forest have also been successively explored to perform classification of geological domains based on sample geochemical information (Cevik et al., 2019). However, not many applications of ML have been employed in measuring blast movement. The ML algorithms being discussed in the following paragraphs are novel and sets the tone for further research.

In the first scenario, three new hybrid models of Support Vector Machines (SVR); a genetic algorithm (GA), an artificial bee colony algorithm (ABC), a cuckoo search algorithm (CS), abbreviated as the GA-SVR, ABC-SVR and CS-SVR respectively, were proposed for the prediction of rock movement in the Husab Uranium Mine in Namibia, the Coeur Rochester Mine, USA and the Phoenix Mine, USA. Eight blasting parameters were used as input variables to develop the model: rock type, number of free faces, first centerline distance, hole diameter, power factor, spacing, subdrill and initial depth of monitoring, and horizontal blast-induced rock movement was the output variable. The use of the hybrid algorithms aided in finding optimal hyperparameters for the final model: i.e *gamma (γ)* and the *penalty factor (C)*. The best performing model was selected by examining the three models. Data collected for all algorithms were divided into training and testing for validation and comparison. The GA-SVR model was designed by

simulating the biological process of evolution where the adaptive abilities of organisms were employed to generate a group of well adapted individuals after continued evolution. The behavior of scout honeybees in finding food sources close to the hive, inspired the development of ABC-SVR model. Finally, the CS-SVR was inspired by how the cuckoo bird searches, lays and hatches its eggs in the nest of another bird considered as the host bird. During calculation, a Levy flight method is used in the search for new nests in the CS algorithm. Figure 6 shows the framework of the proposed models. Details of this work can be found in (Yu, Shi, Zhou, Rao, et al., 2019).

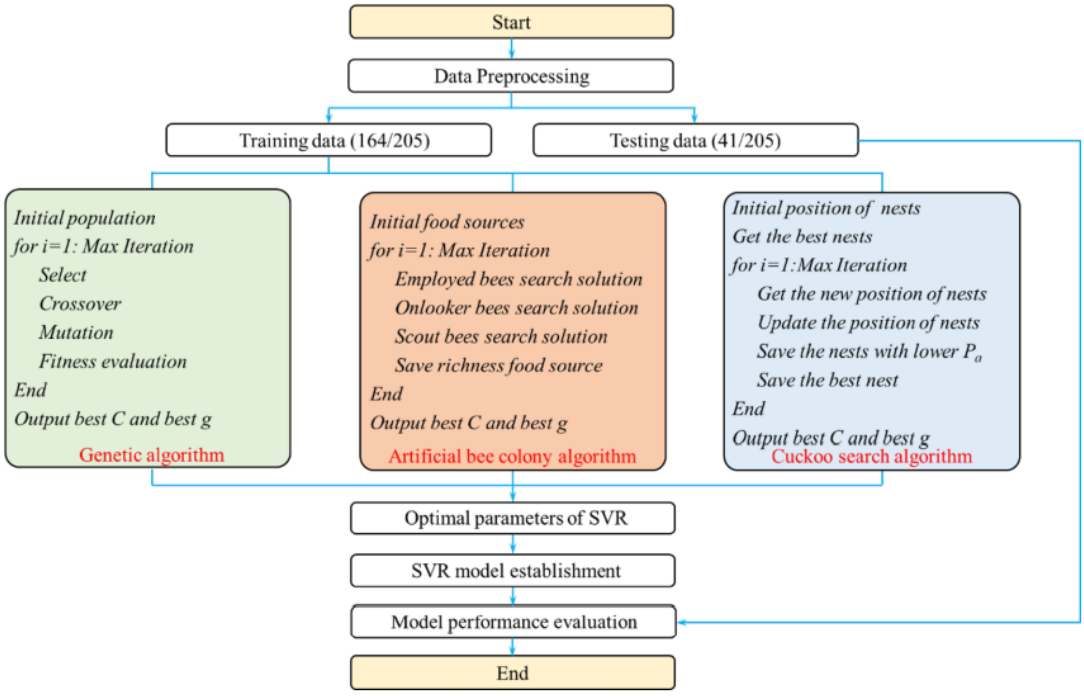


Figure 6 Model framework of GA-SVR, ABC-SVR and CS-SVR *Source: (Yu, Shi, Zhou, Rao, et al., 2019)*

In the second case, three original machine learning techniques: support vector regression (SVR), the Gaussian process (GP), and the extreme learning machine (ELM) were used to develop a predictive model for blast movement. The genetic algorithm (GA) and a whale optimization algorithm (WOA) was used in place of the trial-and-error method, to obtain the optimal hyperparameter search. The ELM, based on neural network theory, was used for its fast-learning ability and good generative performance. The only hyperparameter tuned was the number of neurons. Having extended support vector machine (SVM) from just solving classification but to also solve regression problems, SVR was used. Hyperparameters were γ and *C* as mentioned in section above. GP is a nonparametric model based on random parameters in a gaussian distribution. The mean and covariance functions make up the hyperparameter. The metaheuristic algorithms used were also inspired by natural phenomena just as the previous case. The GA algorithm used is like the one described in the above scenario, and WOA algorithm is designed from the predating nature of whales in the ocean. Like the other swarm-based algorithms, mathematical models inspired by these phenomena are utilized to reduce the error between the predicted values and real values, and the process is terminated when the set error level is reached. Details of this work can be found in (Yu, Shi, Zhou, Gou, et al., 2021).

4. Discussion

The results presented here are case studies from machine learning approaches as proposed by (Yu, Shi, Zhou, Rao, et al., 2019) and (Yu, Shi, Zhou, Gou, et al., 2021). According to literature, BMM methods, even though very costly, provide far better measurement results than the use of visual markers and numerical simulation. However, there is no performance metric in literature that compares their performances with real time or synthetic data.

For the first case study involving GA-SVR, ABC-SVR and CS-SVR, their results together with an artificial neural network (ANN) model were evaluated using correlation coefficient (R^2), mean square error (MSE), variance account for (VAF) and the computing time. Based on the results from these performance metrics, a ranking method was used to the model performance and results are summarized in table 1. From the results, GA-SVR was found to be the best predictive blast movement model and has a faster computing speed.

Table 1 Performance of models Source: (Yu, Shi, Zhou, Rao, et al., 2019)

Method	Model	Results				Rank value				Total rank
		R^2	MSE	VAF	Run time (s)	R^2	MSE	VAF	Run time (s)	
GA-SVR	Training	0.9489	0.0025	94.858	40.56	2	3	2	3	22
	Testing	0.9245	0.0031	91.405		4	4	4		
ABC-SVR	Training	0.9494	0.0024	94.903	48.67	3	2	3	2	20
	Testing	0.9240	0.0031	91.366		3	4	3		
CS-SVR	Training	0.9497	0.0024	94.936	97.28	4	2	4	1	19
	Testing	0.9233	0.0031	91.2842		2	4	2		
ANN	Training	0.8835	0.0237	88.167	2.36	1	4	1	4	15
	Testing	0.9002	0.0172	87.666		1	3	1		

Similarly, in the second case study, the three original ML methods (SVR, GP and ELM), together with two hybrid models each of their kind (GA-SVR, WOA-SVR, GA-GP, WOA-GP, GA-ELM, and WOA-ELM) were evaluated and the best predictive model was selected using a simple ranking method as in the first case study. Results show that WOA-GP obtained the best rank of 53 among the nine models as summarized in table 2. The actual and predicted values of the chosen model is also shown in figure 7.

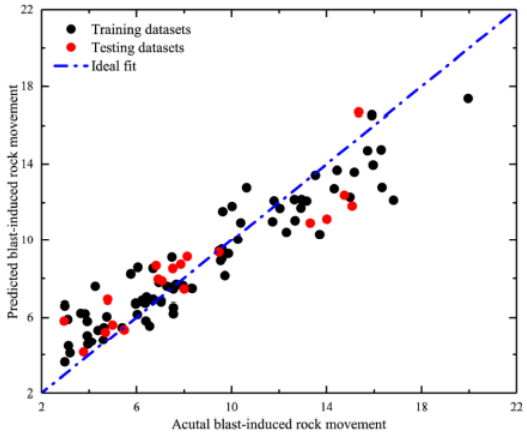


Figure 7 Actual rock movement measurement vrs predicted values by WOA-GP. Source: (Yu, Shi, Zhou, Gou, et al., 2021)

Table 2 Performance of models Source: (Yu, Shi, Zhou, Gou, et al., 2021)

Method	Model	Results			Rank value			Total rank
		R^2	MSE	VAF	R^2	MSE	VAF	
GP	Train	0.828	1.705	82.822	7	7	7	42
	Test	0.788	1.817	78.924	7	7	7	
WOA-GP	Train	0.858	1.475	85.833	9	9	9	53
	Test	0.819	1.679	82.007	9	9	8	
GA-GP	Train	0.858	1.550	85.792	9	8	8	52
	Test	0.819	1.679	82.010	9	9	9	
SVR	Train	0.692	2.284	69.154	1	1	1	6
	Test	0.640	2.367	64.266	1	1	1	
WOA-SVR	Train	0.795	1.862	79.523	4	4	4	18
	Test	0.713	2.116	72.078	2	2	2	
GA-SVR	Train	0.795	1.860	79.564	5	5	5	25
	Test	0.713	2.113	72.160	3	3	4	
ELM	Train	0.725	2.156	72.529	2	2	2	17
	Test	0.719	2.093	72.144	4	4	3	
WOA-ELM	Train	0.800	1.839	80.007	6	6	6	36
	Test	0.781	1.849	78.061	6	6	6	
GA-ELM	Train	0.786	1.904	78.560	3	3	3	24
	Test	0.775	1.871	77.514	5	5	5	

5. Conclusion

Measuring the blast-induced rock displacement is very crucial to reducing ore loss and dilution. Theoretical review has shown that the blast monitoring device (BMM) is very effective in providing a near accurate and reliable measurement of rock displacement than the use of visual markers and numerical modeling. Numerical simulation models have their own merits but the challenges to be addressed to provide a more accurate model persists. There has been considerable amount of research in this area, nonetheless. The machine learning approaches discussed have also proven been effectual in predicting material movement, reducing misclassification, and subsequently providing dig limits for shovels and reducing losses. The data collected from direct measurement was almost consistent with the predicted values of the ML model. Even though the approach is novel, it sets the tone for further exploration of its use for blast movement monitoring.

6. References

Abuntori, C. A., Al-Hassan, S., & Mireku-Gyimah, D. (2021). Assessment of Ore Grade Estimation Methods for Structurally Controlled Vein Deposits—A Review. Ghana Mining Journal, 21(1), 31–44. https://doi.org/10.4314/gm.v21i1.4

Adam, M., & Thornton, D. (2004). A New Technology for Measuring Blast Movement. 6.

Blair, D., & Minchinton, A. (1997). On the damage zone surrounding a single blasthole. Fragblast, 1(1), 59–72. https://doi.org/10.1080/13855149709408390

Cevik, S. I., Olivo, G., & Ortiz, J. M. (2019). Knowledge discovery from geochemical data with supervised and unsupervised methods (Annual Report 2019 No. 2019–02; Predictive Geometallurgy and Geostatistics Lab, pp. 17–29). Queen's University.

Cundall, P. A. (1980). UDEC - A Generalised Distinct Element Program for Modelling Jointed Rock. (PCAR-I-80). Peter Cundall Associates, European Research Office. https://apps.dtic.mil/sti/pdfs/ADA087610.pdf

Fitzgerald, M., York, S., Cooke, D., & Thornton, D. (2011). Blast Monitoring and Blast Translation – Case Study of a Grade Improvement Project at the Fimiston Pit, Kalgoorlie, Western Australia. NEW ZEALAND, 13.

Furtney, J., Sellers, E., & Onederra, I. (2013). Simple models for the complex process of rock blasting. 275–282. https://doi.org/10.1201/b13759-36

Hustrulid, W. A. (2011). Blasting. Encyclopedia Britannica. https://www.britannica.com/technology/blasting

Lawrence, R. W. (1944). Mechanism of Detonation iin Explosives. A Journal of General and Applied Geophysis, 9(1). https://pubs.geoscienceworld.org/geophysics/article/9/1/1/74049/Mechanism-of-detonation-in-explosives

Persson, P.-A. (1997). The relationship between strain energy, rock damage, fragmentation, and throw in rock blasting. Fragblast, 1(1), 99–110. https://doi.org/10.1080/13855149709408392

Preece, D. S., Tidman, J. P., & Chung, S. H. (1997). Expanded Rock Blast Modeling Capabilities of DMC-BLAST, Including Buffer Blasting (No. DEAC04-94AL8500). Sandia National Laboratories, IC1 Explosives Canada. https://digital.library.unt.edu/ark:/67531/metadc685897/m2/1/high_res_d/432902.pdf

Rosa, D. L., & Thornton, D. (2011). Blast Movement Modelling and Measurement. 13.

Schamaun, J. T. (1986). Methods for predicting rubble motion during blasting. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, 23(3), 100. https://doi.org/10.1016/0148-9062(86)91162-9

Shen, C., Laloy, E., Albert, A., Chang, F.-J., Elshorbagy, A., Ganguly, S., Hsu, K., Kifer, D., Fang, Z., Fang, K., Li, D., Li, X., & Tsai, W.-P. (2018). HESS Opinions: Deep learning as a promising avenue toward knowledge discovery in water sciences [Preprint]. Catchment hydrology/Modelling approaches. https://doi.org/10.5194/hess-2018-168

Sołtys, A., Twardosz, M., & Winzer, J. (2017). Control and documentation studies of the impact of blasting on buildings in the surroundings of open pit mines. Journal of Sustainable Mining, 16(4), 179–188. https://doi.org/10.1016/j.jsm.2017.12.004

Taylor, S. L. (1995). Blast induced movement and its effect on grade dilution at the Coeur Rochester Mine [Master's].

Thornton, D. M. (2009). The Application of Electronic Monitors to Understand Blast Movement Dynamics and Improve Blast Designs. 14.

Thornton, D., Sprott, D., & Brunton, I. (2005). Measuring Blast Movement to Reduce Ore Loss And Dilution. https://blastmovement.com/wp-content/uploads/2005/09/2005_ISEE.pdf

Vasylchuk, Y. V. (2019). Advanced Grade Control with Multivariate Geostatistics, Blast Movement Modeling, and Optimized Dig Limits. 236.

Vasylchuk, Y. V., & Deutsch, C. V. (2018). Improved grade control in open pit mines. Mining Technology, 127(2), 84–91. https://doi.org/10.1080/14749009.2017.1363991

Vasylchuk, Y. V., & Deutsch, C. V. (2019). Approximate blast movement modelling for improved grade control. Mining Technology, 128(3), 152–161. https://doi.org/10.1080/25726668.2019.1583843

Yu, Z., Shi, X., Zhou, J., Gou, Y., Rao, D., & Huo, X. (2021). Machine-Learning-Aided Determination of Post-blast Ore Boundary for Controlling Ore Loss and Dilution. Natural Resources Research, 30(6), 4063–4078. https://doi.org/10.1007/s11053-021-09914-5

Yu, Z., Shi, X., Zhou, J., Rao, D., Chen, X., Dong, W., Miao, X., & Ipangelwa, T. (2019). Feasibility of the indirect determination of blast-induced rock movement based on three new hybrid intelligent models. Engineering with Computers, 37(2), 991–1006. https://doi.org/10.1007/s00366-019-00868-0

Zhang, S. (1994). Rock movemnt due to blasting and its impact on ore grade control in Nevada open pit gold mines [Master's]. University of Nevada.

Zhang, Z.-X. (2016). Theory of Detonation. In Rock Fracture and Blasting (pp. 197–216). https://doi.org/10.1016/B978-0-12-802688-5.00009-9

Zou, D. (2017). Contour Blasting Technique for Surface Excavation. In D. Zou, Theory and Technology of Rock Excavation for Civil Engineering (pp. 325–342). Springer Singapore. https://doi.org/10.1007/978-981-10-1989-0_10

Data driven approaches for estimating bulk ore sorting value[1]

Fouad Faraj (ffaraj@minesense.com)
Julian M. Ortiz (julian.ortiz@queensu.ca)
Jose Arnal (jose.arnal@mail.utoronto.ca)

Abstract

ShovelSense is a robust shovel mounted X-ray fluorescence sensor that can measure multiple element grades of each bucket as it is dug at the mine face. The shovel sensor system allows for the bulk sorting of ore and waste at the mine face ensuring each truck is sent to its correct destination. Due to the previous inability of sorting at the truck scale there are not many established methods for predicting the bulk ore sorting value at a given deposit. The ShovelSense grade of the truck load is used to define its final destination, which can be different from the one defined by the short term plan, which is done at block resolution. In this paper, we present two approaches for quantifying the value of the truck reassignments based on the measured grade. First, truck loads within the block are assumed to follow a simple gamma distribution. The second method uses geostatistical simulation at point support to average the grades at truck or block resolution. Both distinct data driven methods predict the potential bulk ore sorting value based on the mine's current operating selectivity and natural variability drawn from blastholes or the short term block model. The bulk ore sorting value predictions are validated with ShovelSense truck diversions from a dispatch dataset of 28,418 trucks at a low-grade, high-tonnage homogeneous Cu porphyry deposit. In addition to the algorithms and workflows presented here, recommendations based on the potential and limitations of each method are given to practicioners seeking to evaluate the bulk ore sorting opportunity for any open pit operation.

1. Introduction

The natural variability linked to the mineralization of ore deposits and mining operational complexity make ore control challenging, resulting in the inevitable loss of ore and dilution of waste in the ore stream. The accurate sorting of material especially at ore-waste contacts is a significant challenge in the mining industry which scales with the deposit's variability (Amirá et al., 2019) and poor grade control practices (Vasylchuk and Deutsch, 2018). Routine grade control relies heavily on the measurement of element grades for samples from boreholes, estimation of blast movement, and constant monitoring with geologists at the mine face. Blastholes are drilled and sampled on a grid with a spacing that can range from 3 to 10 m depending on the material being blasted. These grades are used to estimate grades into a block model from which a dig plan is generated.

[1]Cite as: Faraj F, Ortiz JM, Arnal J (2022) Data driven approaches for estimating bulk ore sorting value, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-07, 96-118.

The sparsity of grade measurements, the difficulty of accurately sampling a blasthole, smoothing introduced through estimation, and material displacement due to blast movement severely limit the performance of the current grade control processes (Rossi and Deutsch, 2013). Furthermore, all decisions based on this model are made at block support, which is usually much larger than the truckload support. This leads to the inevitable ore loss to waste and waste diluting the ore stream, reducing the efficiency of downstream processes.

ShovelSense is a shovel mounted robust X-ray fluorescence (XRF) based sensor which can predict elemental grades in each bucket as it is loaded before being dumped into a truck (Figure 1). This allows the mine to selectively exploit grade variations at a resolution that was not available before. Significant value can be achieved through bulk ore sorting by reducing the amount of ore loss to the waste stream and removing waste from material destined for processing. In a typical setup, the fleet management system (FMS) informs ShovelSense by identifying the buckets that were combined in a truck and the classification of that truck from the mine plan (e.g., ore or waste). ShovelSense aggregates the selected ShovelSense bucket grade predictions to the truck of interest and determines the material classification using the predicted grades. The predicted material classification is transmitted back to the FMS. If it is different from the original estimated material classification, the FMS can redirect the truck to the correct destination in a completely automated fashion requiring no action from the shovel operator or dispatcher.

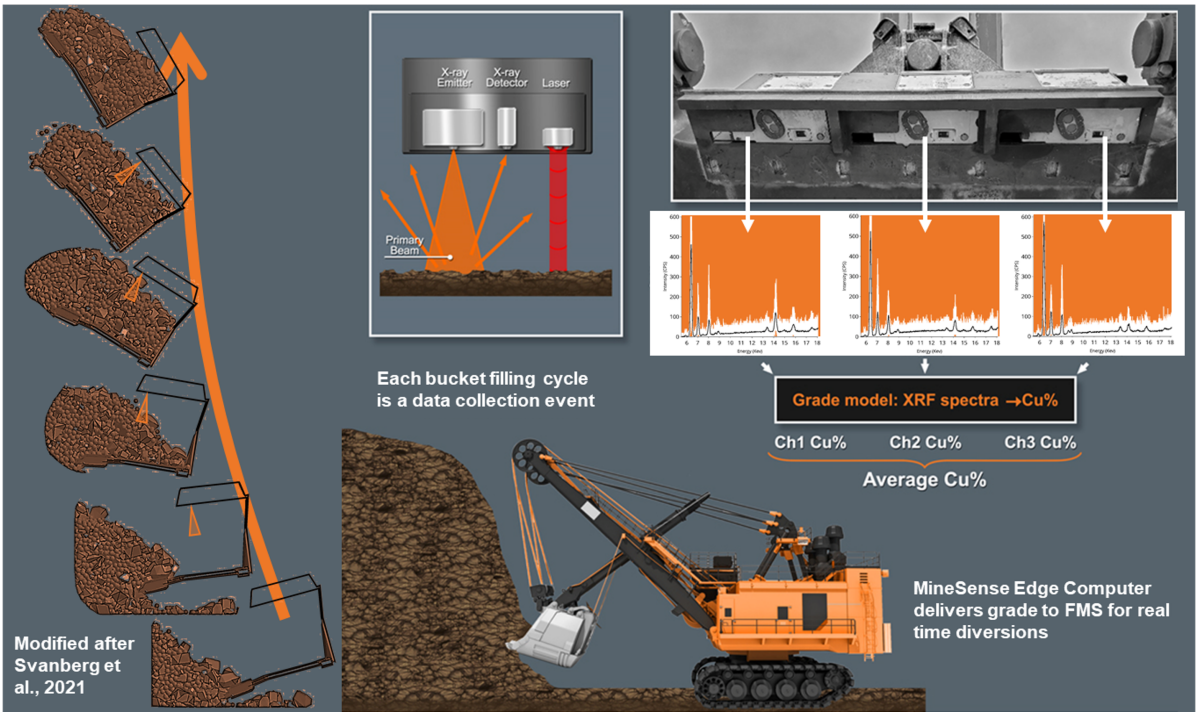


Figure 1: Schematic of the bucket filling and aggregate XRF spectra acquisition. Fill profile modified after Svanberg et al. (2021).

Numerous studies have identified the potential for bulk ore sorting revealing that the ore heterogeneity (Nadolski et al., 2016; Moss et al., 2018) and the mine's current ore control efforts (Sanhueza Passache, 2021) are key drivers in the sorting value in addition to other relevant vari-

ables such as the sorting efficiency, metal price, processing cost, and operation scale (Li et al., 2022). Bulk ore sorting evaluation methods will be critical as more mines consider implementing sensor-based ore sorting systems and decide which are the optimal loading units for sensor installations. Due to the previous inability of sorting at the truck scale, there are currently limited published studies quantifying and validating the potential bulk ore sorting value (Li et al., 2019, 2021). Proposed in this paper are two distinct methodologies for assessing the potential bulk ore sorting value based off truck scaled blasthole grades and variabilities within the selective mining unit (SMU). The validation of the bulk ore sorting predictions is done by comparing the inferred value with the value measured from ShovelSense truck diversions at a Cu porphyry mine.

2. Bulk Ore Sorting Value Prediction Workflows

Both bulk ore sorting value prediction approaches are based off the potential revenue generated from ore recovery and dilution reduction diversions resulting from the reduction of the mine's current SMU down to the truck level. The highest resolution ore control data available is used in a distinct way for each approach to estimate the grades and variability of grades within the current SMU and discretized to truck sized blocks. The two workflows are a simpler gamma distribution approach which is easily automated, and a geostatistical dense simulation approach requiring variogram modeling. Both methodologies start by using or creating an SMU block model with the ore control data available and discretizing it to truck sized blocks. Then the theoretical bulk ore sorting value resulting from the ore recovery and dilution reduction truck block diversions from the SMU block is quantified. Both workflows have the same steps when discretizing to truck sized blocks and quantifying the value but the grade and variability estimation which drives that value is distinct (Figure 2).

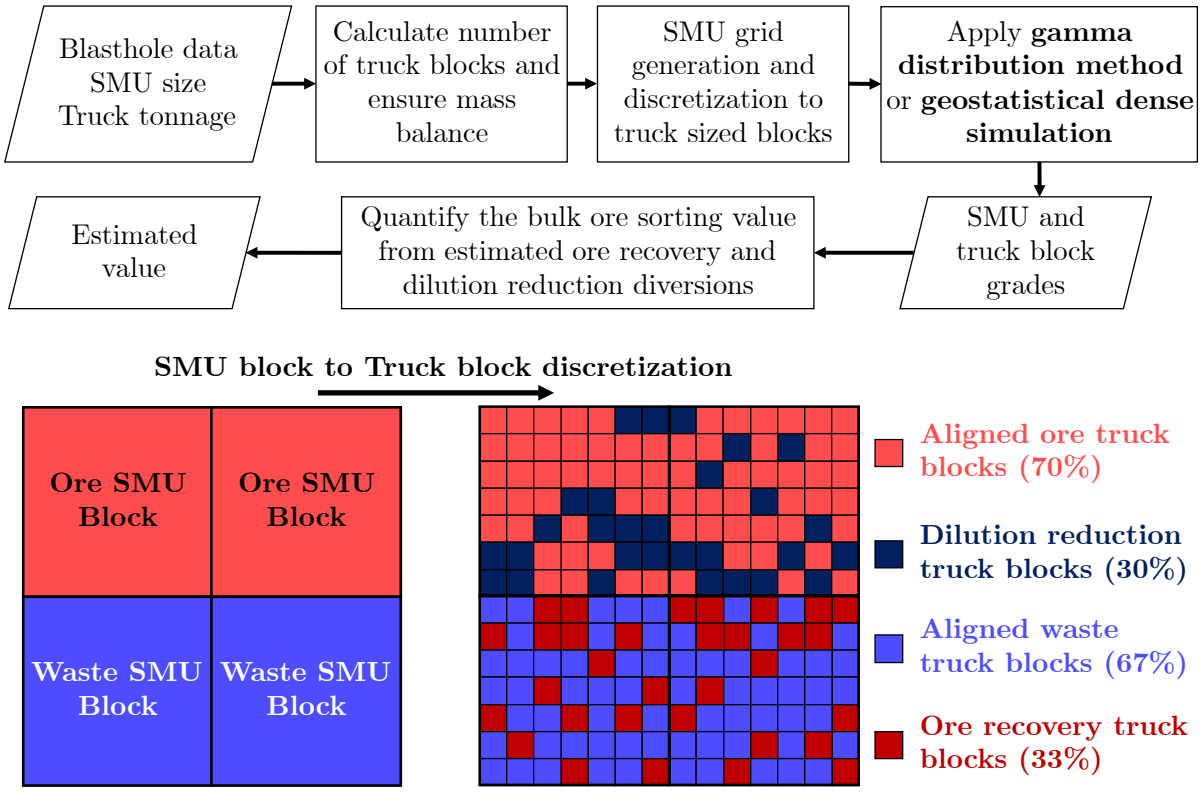


Figure 2: Overall methodology for predicting the bulk ore sorting value with either the simple gamma distribution or geostatistical dense simulation approach.

2.1. SMU Discretization to Truck Sized Blocks

The discretization of the SMU blocks to truck sized blocks needs to achieve a mass balance. The mass of each block is a function of density. For deposits with varying densities this variability needs to be accounted for. More truck blocks are thus assigned to regions of denser material as the tonnage is constrained by the truck type. The formula for calculating the number of truck blocks for each SMU is a simple function of the SMU dimensions, density, and truck capacity which is rounded to the nearest whole number:

$$\mathrm{T}_N = \mathrm{round}\left(\frac{\mathrm{SMU}_l * \mathrm{SMU}_w * \mathrm{SMU}_h * \mathrm{SMU}_\rho}{\mathrm{T}_C}\right), \tag{1}$$

where l, w, h, ρ are the SMU length, width, height, and density respectively and T_C is the truck capacity.

Figure 3 illustrates two extremely distinct cases where a 20x20x15 m SMU block is discretized to truck sized blocks for two different densities and truck capacities. Larger SMUs along with smaller truck capacities will result in more truck blocks for each SMU block and the increased selectivity will better separate ore and waste generating more value.

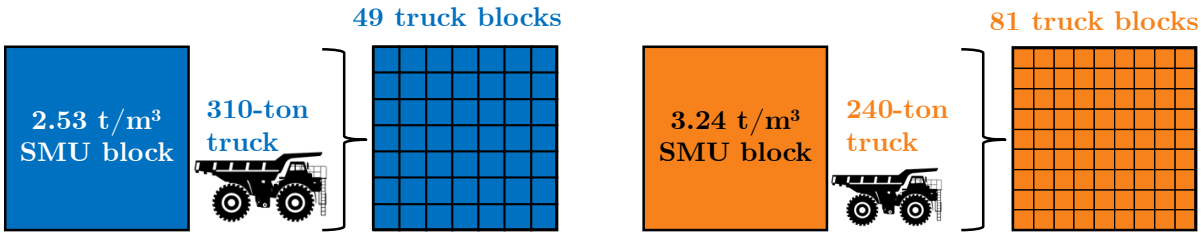


Figure 3: Schematic illustrating the discretization of a 20x20x15 m SMU block to truck blocks for distinct material densities and truck capacities.

When considering multiple payable or deleterious elements used in a net smelter return or similar value based cutoff, only the elements which the bulk ore sensing system can measure should be discretized. In cases where there is an element which cannot be measured by the shovel mounted sensor, the value of that element for all the truck blocks should be set equal to the corresponding SMU block. Generally, the more relevant elements the sensor can measure at a mine, the more valuable the bulk ore sorting tool will be as it provides the highest resolution information currently possible for ore control. Notice that the truckload volume extracted is approximated by a rectangular prism with the SMU height and a smaller area in the XY plane. This is not exactly the geometry of the volume extracted by the shovel to load a truck, but provides a good approximation of the selectivity associated to the truckloads.

2.2. Simple Gamma Distribution Approach

When the blasthole sampling at a given deposit is not preferential, it can be modeled using a theoretical distribution without the need for declustering (Chiles and Delfiner, 2012). In cases where the blasthole sampling is preferential, it must be declustered. Alternatively, the short-term block model can be used as the input source of grades as it is regularly gridded. Geochemical element concentrations of ore grades are always positive, have skewed distributions, and are typically modeled using lognormal distributions (Faraj and Ortiz, 2021) or gamma distributions (Pizarro Munizaga, 2011; Emery, 2012). For the purposes of estimating the Cu grade in a limited number of smaller blocks within an SMU, the gamma distribution is chosen as the longer tail of the lognormal distribution could complicate the SMU reconciliation, especially when the variance is high (Cadigan and Myers, 2001). When varying multiple elements different distributions could be chosen based on which best fit each element, and ideally accounting for their relationships. The probability density for the gamma distribution is given by

$$P(x) = \frac{x^{k-1}\mathrm{e}^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)}, \tag{2}$$

where k is the shape parameter which controls the skewness, θ is the scale parameter which controls the spread of the distribution, and $\Gamma(k)$ is the Euler Gamma function. The shape and scale factor can be written as functions of the mean and variance

$$k = \frac{\sigma^2}{\mu}, \qquad \theta = \frac{\mu^2}{\sigma^2} \tag{3}$$

where μ is the mean and σ^2 is the variance.

The gamma distribution method is simple and based mainly off the blasthole data with another few parameter inputs as detailed in Algorithm 1 and illustrated in Figure 4. Aside from the blasthole data, SMU dimensions, and truck capacity, which are all fixed, the search radius is the only parameter which needs to be defined by the practitioner and could be set to 1.25 times the SMU horizontal length as done here. Squared inverse distance weighted estimates (IDW^2) within the search radius are used to assign the grade of each SMU block and the variance is taken from all the matched blastholes within the radius as well. Using this gamma distribution, the proportion of truckload blocks that are above and below the economic cutoff grade can be determined to quantify the quantity diverted from their original SMU assignment. The exact position within the SMU of these blocks is irrelevant, only their quantity is of significance.

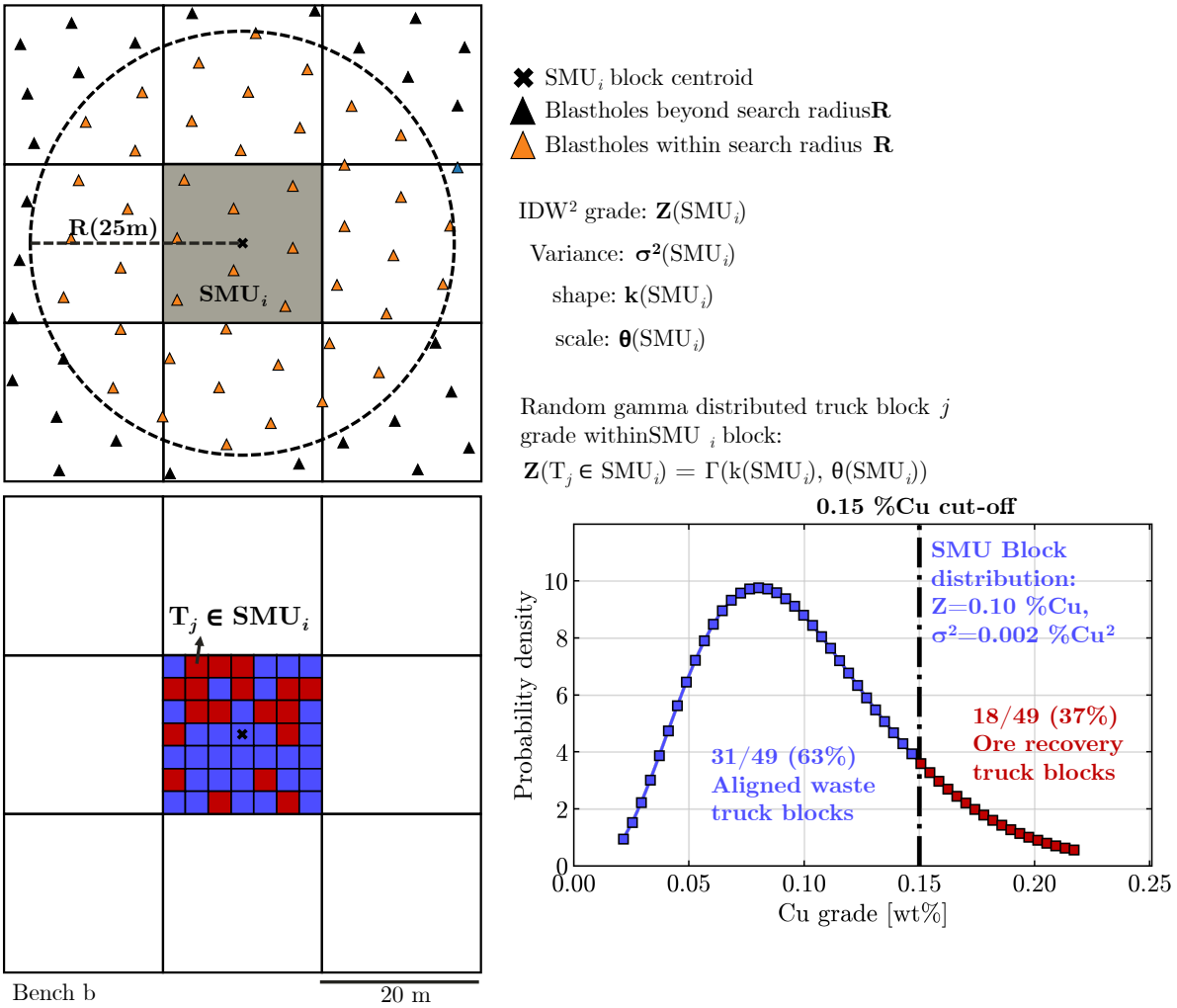


Figure 4: Methodology for assigning SMU and truck block grades with the simple gamma distribution approach with an example for a SMU block with a mean grade of 0.10 %Cu and variance of 0.002 %Cu^2.

Algorithm 1: Gamma Distribution Grade Assignment

Input: Blasthole data (**BH**), SMU data (**SMU**), truck block data (**T**)
Input: Benches (b), Search radius (R), Blasthole positions ($P(\mathbf{BH})$), and grades ($Z(\mathbf{BH})$)
Input: Centroids for each **SMU** block i ($P(\mathbf{SMU}_i)$), and truck blocks j within each **SMU** block i ($T_j \in \mathbf{SMU}_i$)
Output: **SMU** block grades, variance, and gamma distributed grade for each truck block

for *each bench* b **do**

/* Filter blastholes, SMU blocks, and truck blocks to bench b */

$$P(\mathbf{BH}), P(\mathbf{SMU}_i) \in b$$

for *each* ***SMU*** *block* i **do**

/* Search for all matched blastholes ($\mathtt{BH}'$) where the distance d of $P(\mathtt{BH})$ and $P(\mathtt{SMU}_i)$ is within R */

$$\mathbf{BH}' = \{\mathbf{BH} | d(\mathbf{BH} - \mathbf{SMU}_i) \leq R\}$$

/* Assign IDW^2 SMU grade ($Z(\mathtt{SMU}_i)$) and variance ($\sigma^2(\mathtt{SMU}_i)$) from the matched blasthole grades ($Z(\mathtt{BH}')$), average grade ($\overline{Z}(\mathtt{BH}')$), and distances ($d(\mathtt{BH}')$) */

$$Z(\mathbf{SMU}_i) = \sum_{n=1}^{N} \frac{Z(\mathbf{BH}'_n)}{d(\mathbf{BH}'_n)^2} \div \sum_{n=1}^{N} \frac{1}{d(\mathbf{BH}'_n)^2}$$

$$\sigma^2(\mathbf{SMU}_i) = \frac{1}{N} \sum_{n=1}^{N} \left(Z(\mathbf{BH}'_n) - \overline{Z}(\mathbf{BH}'_n) \right)^2$$

/* Calculate the SMU shape ($k(\mathtt{SMU}_i)$) and scale ($\theta(\mathtt{SMU}_i)$) */

$$k(\mathbf{SMU}_i) = \frac{Z(\mathbf{SMU}_i)^2}{\sigma^2(\mathbf{SMU}_i)}, \quad \theta(\mathbf{SMU}_i) = \frac{\sigma^2(\mathbf{SMU}_i)}{Z(\mathbf{SMU}_i)}$$

/* For each of the truck block j draw random gamma distributed grade ($\Gamma(k, \theta)$) and scale the truck block grades ($Z_j(\mathtt{T})$) with the SMU grade ($Z(\mathtt{SMU}_i)$) divided by the average truck block grade ($\overline{Z(\mathtt{T})}$) */

$$Z_j(\mathbf{T}) \sim \Gamma(k(\mathbf{SMU}_i), \theta(\mathbf{SMU}_i)), \quad Z_j(\mathbf{T}) = Z_j(\mathbf{T}) \frac{Z(\mathbf{SMU}_i)}{\overline{Z(\mathbf{T})}}$$

end

end

2.3. Geostatistical Dense Simulation Approach

The use of geostatistical simulations in mining operations is on the rise to quantify heterogeneity and transfer uncertainty and variability into risk for decision making. A major benefit of simulations is the ability to quantify the risk associated with the estimation by assessing the spatial variability (Vann et al., 2002; Rendu, 2002). Here, densely gridded simulations are used to estimate grades at a truck scale for predicting the potential bulk ore sorting value. Densely gridded simulations have previously been used to develop high resolution mining models of mineral grades (Charifo et al., 2013), which have been used for many applications such as informing mining decisions through the mine value chain (Altinpinar et al., 2020).

The geostatistical dense simulation approach requires only the blasthole data, SMU blocks, and discretized truck block definitions. First, domains must be established if there is domaining information available. Then, for each domain, normal score variogram models are developed and used to generate a number of point support realizations (100 sequential gaussian simulations were used in our case) in a dense grid defined by the practitioner such as 1 by 1 m in each bench. The dense grid is averaged to the truck blocks, and then the truck blocks are also averaged to each SMU block as illustrated in Figure 5. The gaussian simulations proposed serve as an adequate baseline for typical homogeneous porphyry deposits. However for more geologically complex deposits which exhibit a high degree of nonlinear features such as veins, channels or folds the simulations could incorporate locally varying anisotropy (Boisvert and Deutsch, 2011) or different continuity for different grade ranges through an indicator approach, since standard gaussian simulations have been shown to miss complex geological structures (Lee et al., 2007).

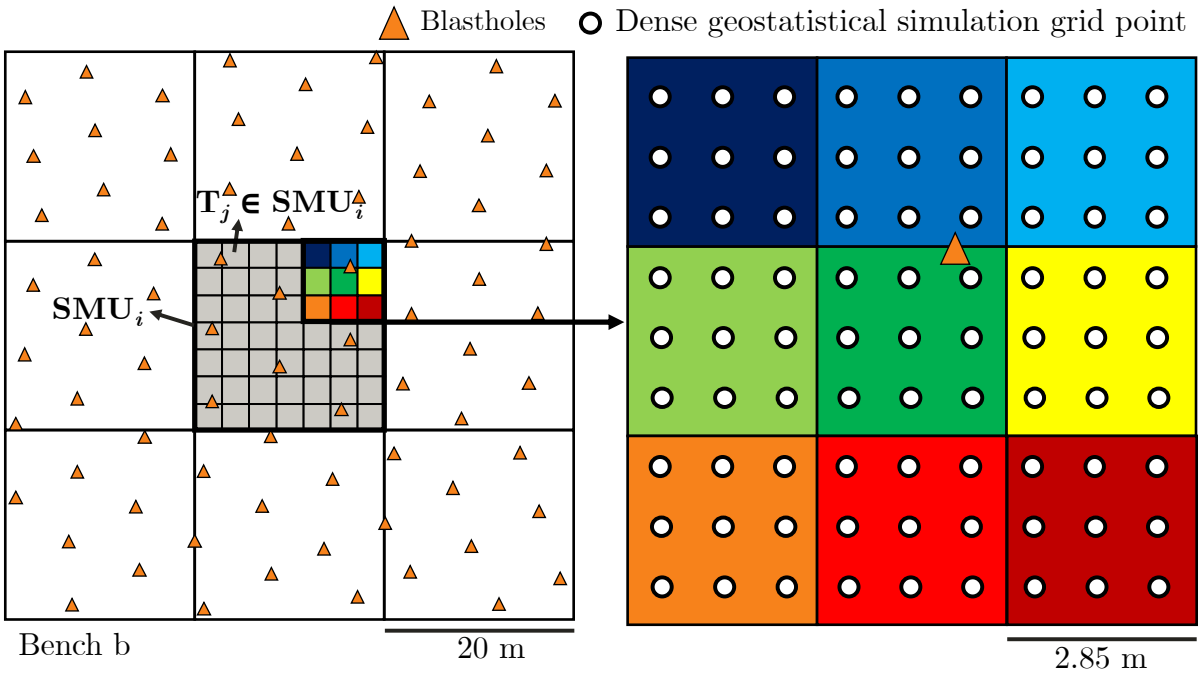


Figure 5: Methodology for assigning SMU and truck block grades with the geostatistical dense simulation approach.

2.4. Value Quantification and Validation Methodology

There are multiple ways to calculate the ore recovery and dilution reduction value from the estimated truck diversions. The optimal way to calculate these will vary from mine to mine. Some mines ideally use a net smelter return to incorporate additional variables and give an accurate evaluation. In contrast others may just work based off the Cu content. We propose a generic method to estimate the ore recovery and dilution reduction value which is applied to both approaches and the dispatch validation data (Figure 6). For the purposes of comparing the predicted value to dispatch data for validation, the assumptions and calculations done to quantify the value are not critical as the same methodology is applied to each. For this deposit, the recovery, recovered value factor, processing cost, and Cu price are taken as 85%, 85%, 5 USD/t, and 3.50 USD/lb respectively.

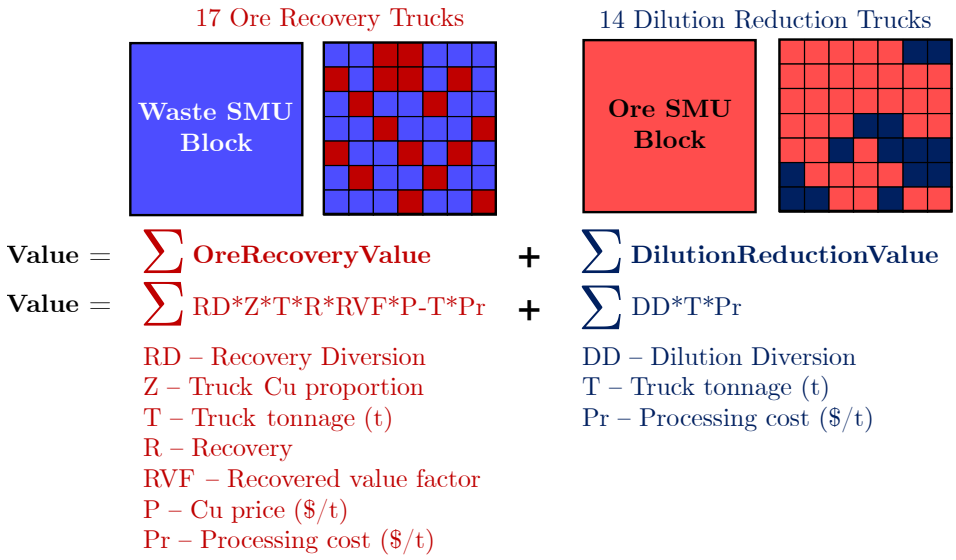


Figure 6: Bulk ore sorting value estimation methodology.

3. Bulk Ore Sorting Value Predictions at a Cu Porphyry Mine

Both the gamma distribution and dense geostatistical simulation approach were applied on a blasthole dataset from a low-grade, high tonnage Cu porphyry mine. The data spans six benches from which ShovelSense truck data was also collected during several months. There are 23,192 blastholes and 28,418 trucks with ShovelSense grades which are used to validate the bulk ore sorting value predictions. Figure 7 shows the blastholes and ShovelSense truck Cu grades throughout the six benches. The Cu porphyry mine has a cutoff grade of 0.15 %Cu and SMU of 20x20x15 m.

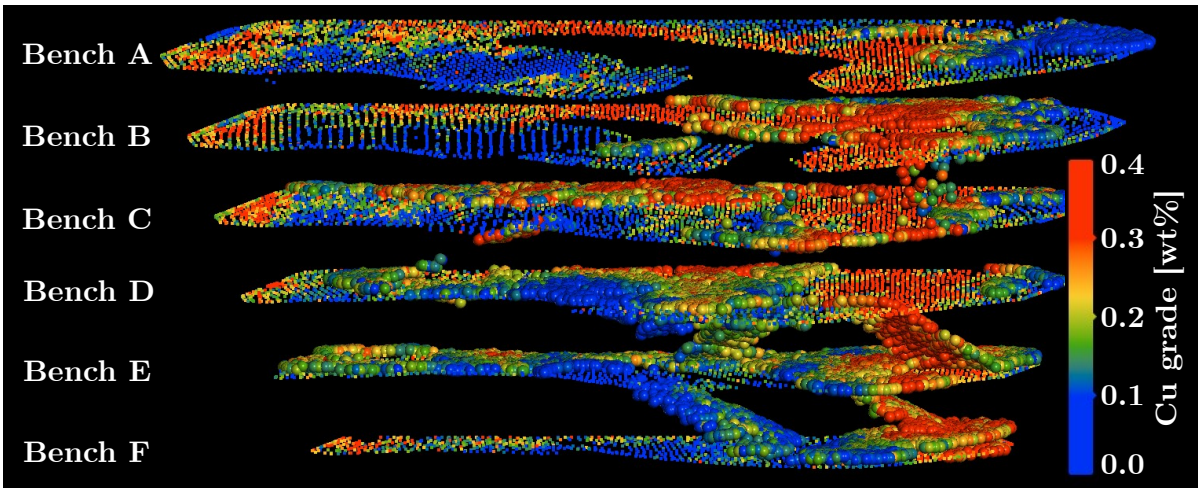


Figure 7: The low-grade, high tonnage Cu porphyry dataset used showing the blastholes and ShovelSense truck grades for each of the six benches. The vertical scale is four times the horizontal scale.

3.1. Discretization Integrity for Tonnage and Metal Balance

With the mine's pit density of 2.54 t/m^3 there will be 49 truck blocks for each SMU block to achieve a tonnage balance as outlined in Table 1. From the 23,192 blasthole data, 2,911 SMU blocks were generated with a corresponding 142,639 truck blocks after filtering out SMU blocks with less than 10 blastholes in the 25 m radius as these represent blocks along the edges without sufficient data to be properly evaluated. In addition to achieving a mass balance, the truck block grades averaged to the corresponding SMU block must also match in metal content. The error metrics tabulated in Table 1 demonstrate that their is no significant metal mismatch. The discretization integrity for tonnage and metal content should always be checked for all considered elements as significant differences could cause errors in the final bulk ore sorting value estimate.

Table 1: Parameters and summary statistics of the SMU to truck block discretization for the tonnage and metal balance.

SMU and Truck Block Tonnage Balance		
Parameter	SMU Block	Truck Block
Width [m]	20	2.857
Length [m]	20	2.857
Height [m]	15	15
Volume [m^3]	6000	122
Density [t/m^3]	2.54	2.54
Tonnage [t]	15,240	310
SMU and Truck Block Metal Balance		
Parameter	GammaDistribution - SMU	GeostatsSimulations - SMU
Count	2911	2911
Min [%Cu]	-9.99e-16	-5.55e-16
Mean [%Cu]	1.46e-18	1.22e-18
Max [%Cu]	8.32e-16	5.55e-16
St Dev [%Cu]	8.53e-17	1.37e-16

For comparing the predicted bulk ore sorting value to the ShovelSense dispatch dataset, the SMU and truck block data was further filtered to only include blocks within 15 m of a dispatched

truck. This filter was applied to improve the comparison by roughly matching the SMU and truck block data to the actual material dug. There will still be some expected noise and errors in the comparison because there were other shovels without ShovelSense installed working in the same area but not far enough to filter out.

3.2. Truck Block Grades and Classification

The truck block grades for the gamma distribution were randomly assigned to each block within the SMU block. The spatial aspect of each truck block is not relevant for the purpose of bulk ore sorting value predictions. The dense simulation accounts for the spatial distribution of the data and the continuity is based on the normal score spherical variogram models developed for the three principal directions of anisotropy given by

$$\gamma(h) = \begin{cases} 0 & \text{for } h = 0 \\ C_0 + C_1 \left[\frac{3}{2}\frac{h}{a} - \frac{1}{2}\left(\frac{h}{a}\right)^3 \right] & \text{for } 0 < h \leq a \\ C_0 + C_1 & \text{for } h > a \end{cases}, \tag{4}$$

where C_0 is the nugget of 0.20, C_1 is the sill contribution of the spherical structure, equal to 0.80, h is the lag distance, a is the range of 250 m, 140 m, and 80 m for the three directions respectively.

The gamma distributed truck block grades show much higher spatial variability than the dense geostatistical simulation which drives the difference in the theoretical ore recovery and dilution reduction from the SMU. A zoomed in section on the SMU block model in Figure 8 highlights the difference in the spatial distribution of grades for each truck block within the SMU blocks. There are even many blastholes with different classification from their host SMU block which is also captured by the gamma distributed grades but not by the dense geostatistical simulation throughout this variable zone. This is explained because, even for the truck blocks, the geostatistical simulation takes into account the change of support, that is, the fact that the truckload is much larger in volume than the blasthole.

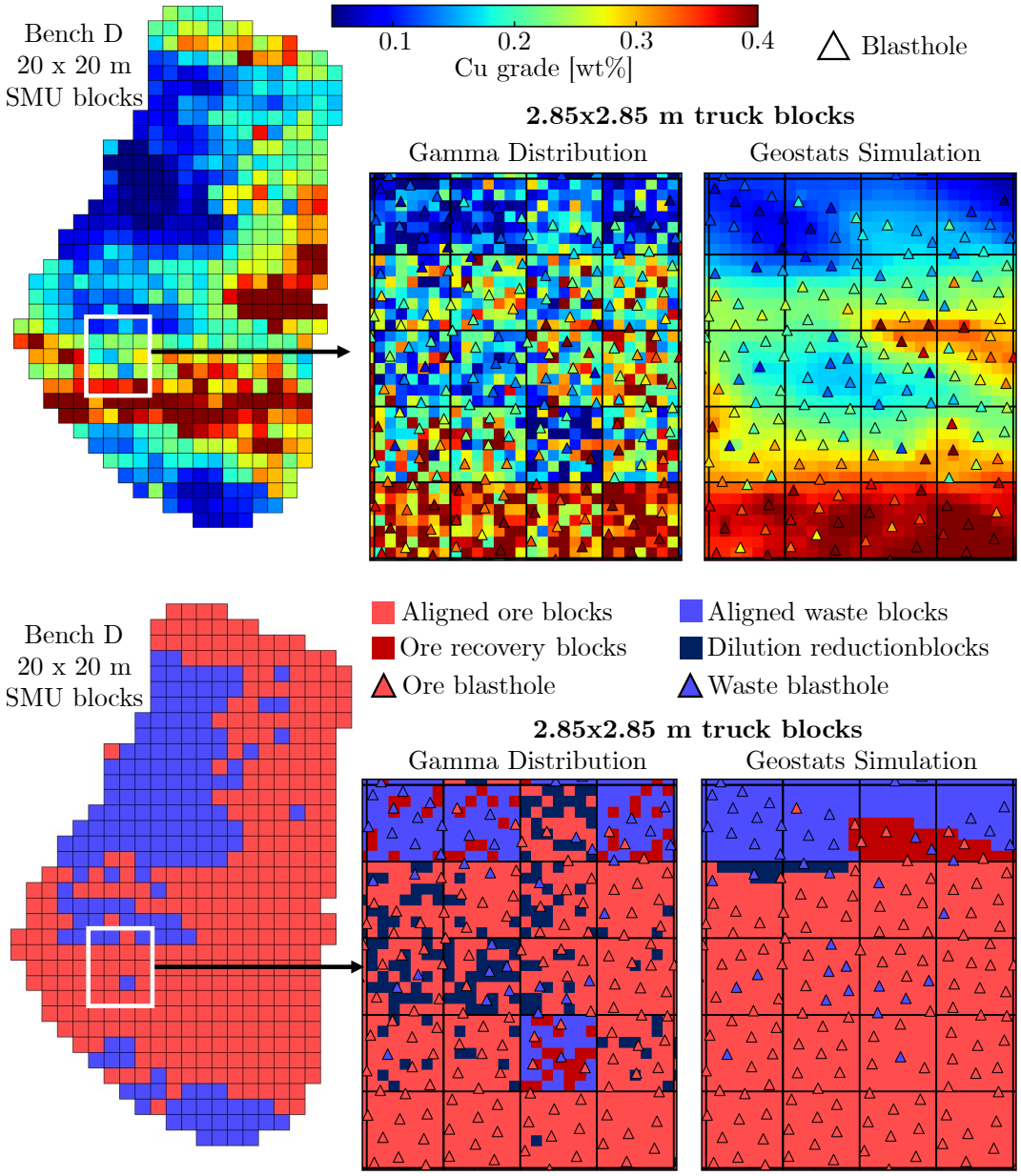


Figure 8: Spatial plots comparing the SMU and truck block grades and classifications for the gamma distribution and geostats simulation within a variable section.

The SMU grade and variability has the biggest influence on the proportion of truck blocks classified differently than the SMU blocks. Figure 9 shows that most of the difference in material classifications occur with more variable SMU blocks near the cut-off of 0.15 %Cu. Even highly variable blocks don't result in many theoretical diversions if the SMU grade is significantly higher or lower than the cut-off. Compared to the ShovelSense truck grades, the gamma distribution was more variable while the geostatistical simulation was less variable (Figure 10). Despite applying the 15 m filter, the ShovelSense data still represents less than half of the predicted tonnages and

could have dug more homogeneous areas.

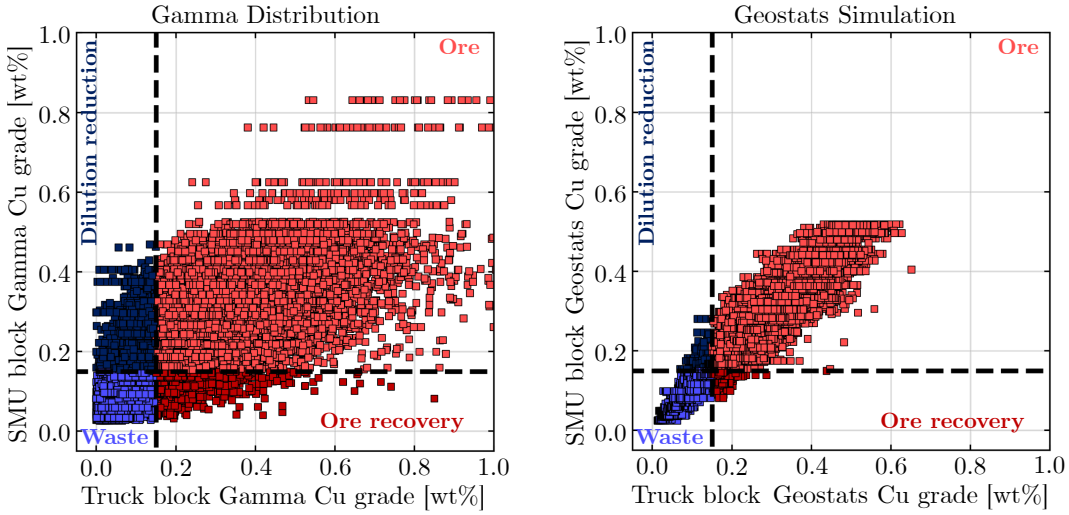


Figure 9: Scatter plots of the truck block grades against their respective SMU block grade for the gamma distribution and geostats dense simulation approaches with a black dashed line highlighting the 0.15 %Cu cut-off.

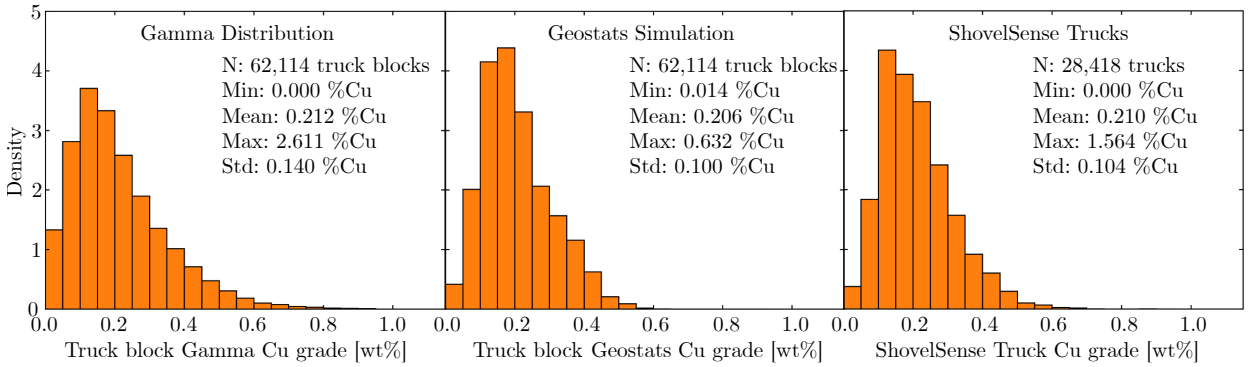


Figure 10: Histograms showing the distribution of Cu grades for the gamma distribution, geostats simulation, and ShovelSense trucks.

The material classification predictions as either ore or waste varied significantly between the predictions and ShovelSense dispatch diversions (Figure 11). The gamma distribution approach predicted a total of 20.1% diversions which is similar to the 22.7% measured by the ShovelSense dispatch data but the predictions resulted in much more dilution reduction than ore recovery which is opposite of the ShovelSense dispatch data. The geostats simulation significantly underpredicted the diversions at only 6.1%. The main discrepancy between the deviation types of the gamma distribution and ShovelSense dispatch data is due to a cut-off change and short term blending campaigns carried out during the study period without the data being available to account for it. Nevertheless the diversions are high due to the majority of the grades being close to the cut-off which is expected of a homogeneous, low-grade, and high-tonnage Cu porphyry deposit.

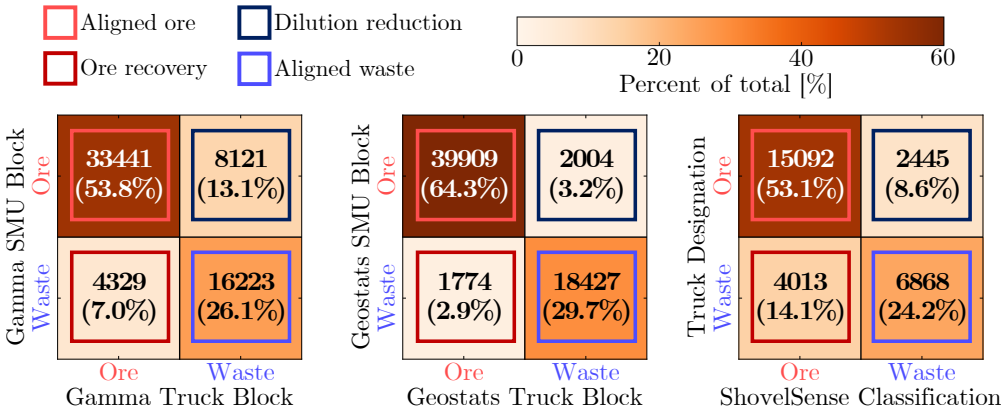


Figure 11: Confusion matrices showing the distinct material classifications of the gamma distribution and geostatistical simulation predictions as well as ShovelSense compared to the dispatch classification.

3.3. Comparing the Predicted and ShovelSense Diversion Value

Since the ShovelSense installed shovel did not dig the entirety of the six benches, there will be predicted volumes not present in the validation data even after applying the filter to only include blocks within 15 m of a ShovelSense truck. In order to make a fair comparison it is important to normalize the predicted value by the tonnage. The weighted average bulk ore sorting value per ton mined for the ShovelSense diversion data is 1.04 USD/t which was best predicted by the gamma distribution at 0.94 USD/t while the geostats simulation predicted a lower value at 0.22 USD/t (Table 2). Interestingly, the gamma method value predictions for benches which did not include data containing ramps lined up with the ShovelSense diversion data by a weighted average of about 9.3% which is 6.6% better than those with ramps at 15.9%. Despite the best efforts some noise such as the effect from the ramps is always to be expected with the comparison of two distinct spatial datasets, especially when considering that the blastholes are fixed in space and the shovel with ShovelSense measures material in situ after having been displaced and mixed by blast movement.

Table 2: Normalized bulk ore sorting value for the two approaches compared to the ShovelSense dispatch data, the average is multiplied to a total using the 8.81 Mt from the 28,418 ShovelSense trucks.

	Bulk Ore Sorting Value per ton mined		
Bench	Gamma Distribution	Geostats Simulation	ShovelSense Dispatch
Bench A [USD/t]	0.80	0.25	0.69
Bench B [USD/t]	0.99	0.22	1.03
Bench C [USD/t]	0.94	0.23	1.09
Bench D* [USD/t]	0.97	0.20	1.13
Bench E* [USD/t]	0.93	0.20	1.2
Bench F* [USD/t]	0.90	0.26	0.83
Weighted average [USD/t]	0.94	0.22	1.04
Total from 8.81 Mt [MUSD]	8.28	1.94	9.16

*Includes a ramp connecting two distinct benches introducing noise in the ShovelSense dispatch data

4. Discussion

4.1. Discrepancies and Alignment in Ore Recovery and Dilution Reduction Estimates

While the value predictions roughly aligned for the gamma distribution approach, there was a 6% average discrepancy in the diversion predictions. The mine labels were assigned based on the mine plan at the time of mining, during which the cutoff changed and there were short term blending campaigns adding inconsistencies to the originally assigned material type compared to the fixed 0.15 %Cu cut-off applied to the ShovelSense truck grades. The effects of changing the cut-off or any kind of special campaigning could be prevented by using the historical ShovelSense classifications at the time of digging but this data was not available. In order to allow for a more fair comparison the gamma distribution method was compared to the ShovelSense diversions based on inverse distance weighted estimates of the blasthole Cu grade to each truck. When compared to the blasthole classifications, the average discrepancy in diversions reduced to an impressive 0.6%. After exluding ramp data, a bench by bench analysis of the diversion predictions with the gamma method correlated to the ShovelSense diversions from blasthole classifications with a pearson correlation coefficient of 0.80 (Figure 12). The weighted average bulk ore sorting value per ton mined estimate based off ShovelSense diversions from the blasthole estimates is 0.90 USD/t which deviates by only 4% from the gamma distribution estimate of 0.94 USD/t.

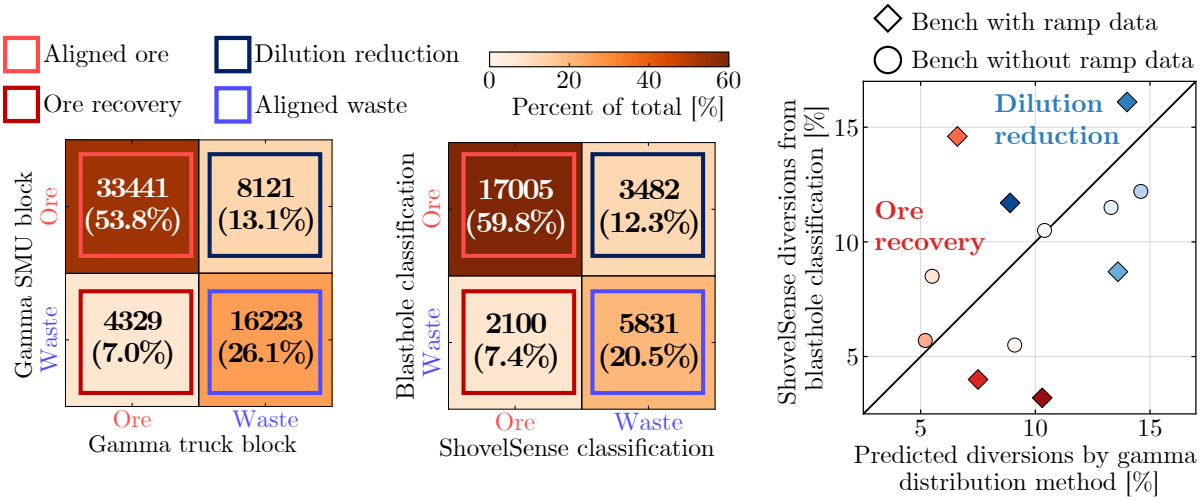


Figure 12: Confusion matrices showing the distinct material classifications of the gamma distribution predictions and ShovelSense compared to a blasthole classification. Also shown a scatter plot of the predicted diversions and the ShovelSense diversions from the blasthole classification for ore recovery and dilution reduction with the different shades representing distinct benches.

In addition to predicting the bulk ore sorting value from diversions, the additional Cu metal and reduced waste processed by the mill can be calculated. The mine plan based on blasthole classifications would have produced 14,607 t Cu metal and generated 2.5 Mt of waste. With ShovelSense the Cu production increases by 9% to 15,990 t Cu metal and generates 2.9 Mt waste which is a 14% reduction in the 0.4 Mt of waste processed by the mill. This compares well to what the mine plan would be based on the gamma distribution method with the SMU blocks normalized to the same 28,418 truck tonnage which would have produced 14,756 t Cu metal and 2.9 Mt waste.

The predicted truck block diversions from the SMU blocks would increase Cu metal by 6% at 15,583 t Cu metal and the waste generated would have been 3.4 Mt waste, an 18% decrease in the 0.5 Mt of waste processed by the mill. The gamma distribution method predictions of Cu metal increase and processed waste reduction align within 4% of the ShovelSense truck diversions from the blasthole classifications.

4.2. Potential and Limitations of the Geostatistical Dense Simulation and Gamma Distribution Approach

The simplicity of the gamma distribution method likely influenced its effectiveness at predicting the bulk ore sorting value over the dense geostats simulation due the little information available for the study. To make the geostatistical simulation approach more robust, domaining information would have been required as treating the entire dataset as a single domain likely had an over smoothing effect (Lee et al., 2007). Despite the simply defined simulation parameters and single domain simulation, the resulting geostats simulation distribution matched the ShovelSense truck grades much better than the gamma distribution. Interestingly the higher variability of the gamma distribution may have influenced why its predicted value better matched the ShovelSense dispatch data since the higher in situ variability replicates the effect of blast movement and mixing which is not accounted for in the static block models. The gamma distribution algorithm could also be fine tuned by capping the max grade or shifting the median grade to match the mean SMU grade depending on the input data for a given deposit.

The theoretical standard deviation at truck support was estimated to be 1.20 %Cu based on the blasthole data using the normal score variogram. Sensitivities were made of the nugget effect to calibrate this truck block variance with the ShovelSense data. The standard deviation for the gamma distribution method was 0.140 %Cu which is 0.020 %Cu higher while ShovelSense was 0.104 %Cu which is 0.016 %Cu lower. Despite the 15 m filter applied, the discrepancy in the variance may have still been influenced by the ShovelSense not capturing more variable areas while the gamma method did evaluate some of the more variable areas where the shovel did not dig. Furthermore the variance in the gamma distribution approach could have been scaled to better fit the deposit. Ultimately the variance of Cu grades depends on the type of deposit (Gerst, 2008) and if the optimal distribution is uncertain, the gamma or lognormal distributions are adequate options for most base metal deposits (Journel, 1980).

When using the proposed, or any bulk ore sorting value estimation tool, it is crucial to consider the data being used. Many mines have shovels which work exclusively in waste where there is no potential for bulk ore sorting and thus using an entire dataset from the mine may lead to inaccurate results. To better predict the value a shovel would unlock with a bulk ore sorting sensor representative data should be used from the pit, phase, or area where it is working. Predictions could even be made for different phases of the mine to determine a sequence for shovel sensor installations based on which would maximize the net present value.

When validating bulk ore sorting predictions as done here there are many variables which are important to consider. The presence of ramps could add noise to the analysis as truck loads by the shovel will need to be assigned to either the bench above or below. Also assuming a fixed cut-off grade could be troublesome if the mine changed their cut-off or had some kind of blending campaigns during the time period being analyzed, these factors could be incorporated if that information is available.

4.3. Influence of Selectivity and Deposit Heterogeneity on the Bulk Ore Sorting Value

The larger SMUs will have a greater bulk ore sorting value when decreasing the selectivity down to the truck scale. The bulk ore sorting benefit correlates with the remnant uncertainty within a mining block which depends on the geological variations and sampling but scales with increasing block size (Chiquini and Deutsch, 2020). Efforts have been made to quantifying the recoverable reserves from exploration drill hole data to the SMU scale (Boisvert et al., 2008) and studies have investigated the benefits and downsides to mining with varying smaller block sizes (Jara et al., 2006). With ShovelSense the selectivity is down to the truck scale and the gamma distribution bulk ore sorting tool can be run with varying SMU sizes and truck capacities to quantify the bulk ore sorting value at a truck scale selectivity.

For the 42.84 Mt mined through six benches in the low-grade, high tonnage Cu porphyry deposit studied here, the gamma distribution bulk ore sorting value prediction tool was run with several distinct SMUs and truck capacities used in the industry for large open pit mines. The SMU was varied from 10x10x15 m to 30x30x15 m and the trucks from 205 t to 400 t. As expected, the most benefit is achieved when reducing a larger SMU to the smallest truck available, however there is no consideration for the significant negative impact the more selective mining would have on production costs. However note that the potential limitation of the current approach is that the support effect is not properly handled. The current data compares sample information taken at point support and averaged over each shovel load with a discretized simulated truckload block. This requires further investigation to ensure the smoothing due to the change of support (from "point" blasthole data to block data) is properly accounted.

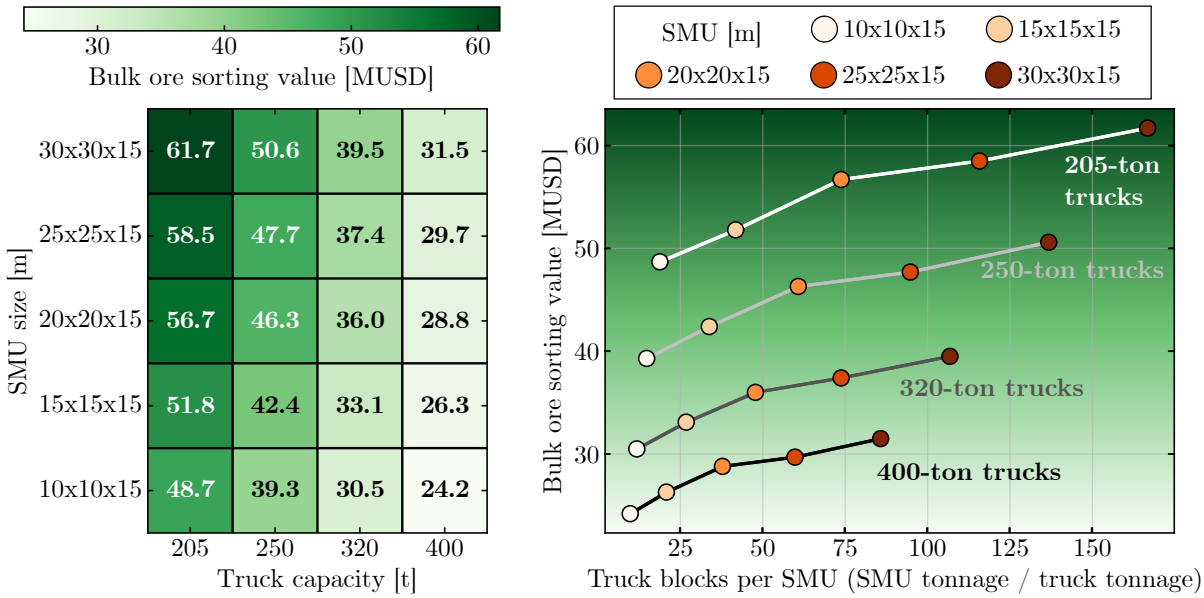


Figure 13: Gamma distribution bulk ore sorting value when digging through 42.84 Mt in six benches varying the SMU and truck capacity to standard values for large open pit mines.

Another important key driver in the bulk ore sorting value is the natural variability of the deposit. The schematic in Figure 14 illustrates how the ore recovery value scales with the SMU

grade variance in four distinct deposits based off the authors' experience. The gamma distribution bulk ore sorting tool from the Cu porphyry data used in this study achieved a strong pearson correlation coefficient of 0.85 when comparing the SMU grade variance to the ore recovery value (Figure 14). The relationship between the grade variability and ore recovery value is expected to scale with more variable deposits. Cu porphyries typically represent some of the more homogeneous deposits which still present significant opportunities for bulk ore sorting when considering the truck scale.

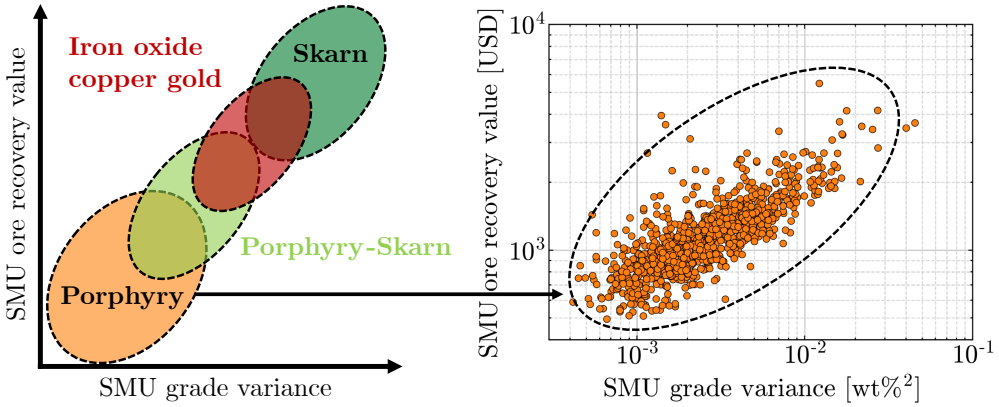


Figure 14: Schematic showing how increasingly heterogeneous deposits generate higher ore recovery value with real data from the gamma distribution bulk ore sorting tool shown for the homogeneous low grade high tonnage Cu porphyry deposit.

As a deposit's heterogeneity increases the mine will benefit more from increasing its selectivity with a bulk ore sorting system. While a system like ShovelSense has a negligible impact on mining, the use of a smaller trucks will likely have a negative impact on production and costs. If the operational complexity can be handled, extremely heterogeneous areas could even justify the use of smaller shovels and bucket level sorting by using two trucks side by side to solely load each one with either ore or waste separately. There is room for optimizing the selectivity and productivity to maximimize the net present value as illustrated in Figure 15. Areas with lower variability will benefit more from a more productive mining method while areas with higher variability will generate additional value with an increased selectivity.

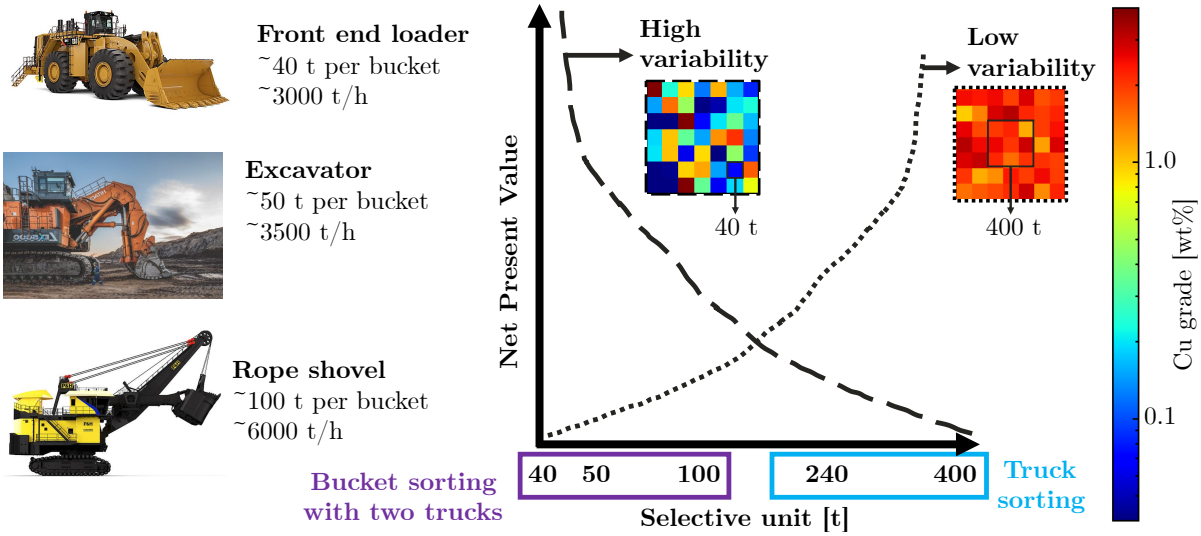


Figure 15: Schematic representing how the maximum net present value is achieved by optimizing the production rate and selectivity based on the loading and haulage systems available. Front end loader picture from Caterpillar (2022), excavator picture from Hitachi (2022), and rope shovel picture from Komatsu (2019).

The gamma distribution bulk ore sorting value prediction tool could even be used by mines which already have ShovelSense installed. Assuming the mine has mining equipment with distinct selectivity and productivity, the predicted bulk ore sorting value can be used to optimize where to send each available unit in the fleet. For example considering the bench shown in Figure 16 more selective loading and hauling units can be sent to dark green areas with high predicted sorting value while bigger, less selective units are sent to more homogeneous areas where the sorting value is predicted to be low as these would benefit more from the increased production rate.

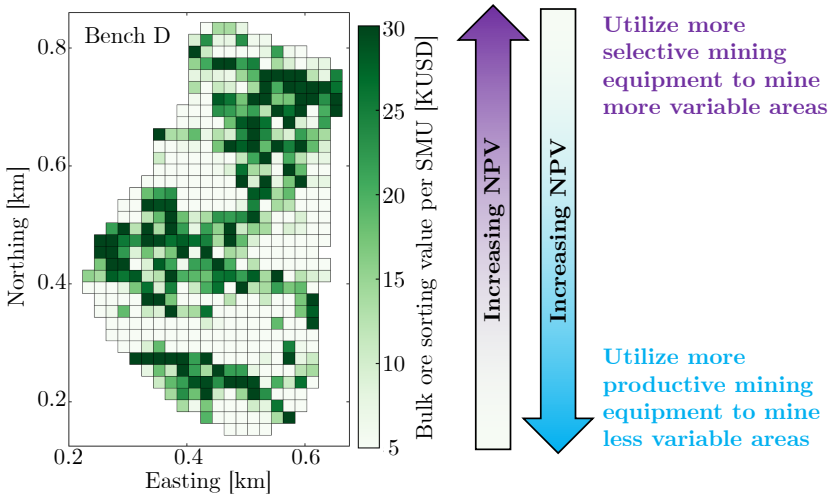


Figure 16: Map showing the bulk ore sorting value prediction for a bench indicating areas which would benefit from either more selectivity or a higher productivity.

5. Conclusion

The discretization of SMU blocks to theoretical truck blocks was proven to be an adequate method for simulating truck grades based on the average grade and variance of blastholes surrounding the SMU blocks. Various bulk ore sorting value parameters based on ShovelSense diversions from their blasthole classification were successfully predicted by the simple gamma distibution method (Table 3). The minor discrepancies between the gamma distribution method predictions and ShovelSense truck data is likely due to a slight overprediction of the variance due to the high degree of homogeneity within the studied deposit. The dense geostats simulation did not compare well with the ShovelSense truck data likely due to poor domain definition and likely a discrepancy in assessing the support effect. A better understanding of the post blast grade variability at the truck scale could also serve to inform and improve the bulk ore sorting value predictions.

Table 3: Summary of the various gamma distribution method predicted parameters compared to the ShovelSense diversions from blasthole classifications.

Parameter	Gamma Distribution Method	ShovelSense/Blasthole diversions	Difference
Ore recovery diversions	7.0%	7.4%	-0.4%
Dilution reduction diversions	13.1%	12.3%	0.8%
Cu metal increase	6%	9%	-3%
Dilution decrease	18%	14%	4%
Sorting value from 8.81 Mt	8.28 MUSD	7.93 MUSD	0.35 MUSD

Here the bulk ore sorting value was estimated for a low-grade, high tonnage Cu porphyry mine. Future work should consider comparing deposits with distinct natural variability as the deposit's heterogeneity likely has the strongest impact on the bulk ore sorting value. While increasingly variable deposits will reap more benefits from bulk ore sorting, even the more homogeneous Cu porphyry mines as the one discussed here present various opportunities to reroute trucks to their correct destinations. Ultimately, when working with a selectivity at the truck scale most magmatic base metal deposits will present various opportunities for bulk ore sorting.

The value presented here only include the immediate benefits from recovering additional ore and reducing dilution. The calculation of these immediate benefits will vary from mine to mine but the most precise information available should be used to get most accurate estimates. There is still a big need to better understand the impact bulk ore sorting has downstream. A few examples include increasing the grade and reducing the variability in the mill feed which could improve recovery (Kurth, 2021), the environmental benefits reducing the tons of CO_2 emitted per ton of concentrate (Sturla-Zerene et al., 2020), or the savings on the operational costs such as pneumatic maintenance, fuel, or electricity on transportation equipment (de Werk et al., 2017), and the liner or steel balls of the SAG mills for comminution equipment (Avalos et al., 2020; Yahyaei et al., 2009).

6. Acknowledgements

MineSense Technologies Limited is acknowledged for supporting this research. Stewart Langley and Miguel Carrera are also acknowledged for their insights and discussions which helped develop the gamma distribution approach presented in this paper.

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2017-04200 and RGPAS-2017-507956.

Contributions to this work were as follows: F. Faraj developed the data analysis and methodological proposal, J. Ortiz supported the theoretical methodology and J. Arnal provided support developing the algorithms.

7. Bibliography

Altinpinar, M., Sari, Y.A., Ortiz, J.M., 2020. Synthetic high-resolution ore deposit model and mine plan. Predictive Geometallurgy and Geostatistics Lab 2020-09, 147–169.

Amirá, R., Morales, N., Cáceres, A., 2019. Analysis of the impact of the dilution on the planning of open-pit mines for highly structural veined-shaped bodies, in: Proceedings of the 27th International Symposium on Mine Planning and Equipment Selection-MPES 2018, Springer. pp. 187–196.

Avalos, S., Kracht, W., Ortiz, J.M., 2020. Machine learning and deep learning methods in mining operations: A data-driven sag mill energy consumption prediction application. Mining, Metallurgy & Exploration 37, 1197–1212.

Boisvert, J.B., Deutsch, C.V., 2011. Programs for kriging and sequential gaussian simulation with locally varying anisotropy using non-euclidean distances. Computers & Geosciences 37, 495–510.

Boisvert, J.B., Ortiz, J.M., Deutsch, C.V., 2008. Local recoverable reserves prediction with block lu simulation. International Journal of Mining and Mineral Engineering 1, 3–21.

Cadigan, N.G., Myers, R.A., 2001. A comparison of gamma and lognormal maximum likelihood estimators in a sequential population analysis. Canadian Journal of Fisheries and Aquatic Sciences 58, 560–567.

Caterpillar, 2022. Large wheel loaders 994k. `https://www.cat.com/en_US/products/new/equipment/wheel-loaders/large-wheel-loaders/1000001147.html`. Accessed: 2022-09-11.

Charifo, G., Almeida, J.A., Ferreira, A., 2013. Managing borehole samples of unequal lengths to construct a high-resolution mining model of mineral grades zoned by geological units. Journal of Geochemical Exploration 132, 209–223.

Chiles, J., Delfiner, P., 2012. Geostatistics: Modeling Spatial Uncertainty. 2-nd ed.-NewYork, Toronto,-JohnWiley'Sons. Inc.

Chiquini, A., Deutsch, C.V., 2020. Mineral resources evaluation with mining selectivity and information effect. Mining, Metallurgy & Exploration 37, 965–979.

Emery, X., 2012. Co-simulating total and soluble copper grades in an oxide ore deposit. Mathematical Geosciences 44, 27–46.

Faraj, F., Ortiz, J.M., 2021. A simple unsupervised classification workflow for defining geological domains using multivariate data. Mining, Metallurgy & Exploration 38, 1609–1623.

Gerst, M.D., 2008. Revisiting the cumulative grade-tonnage relationship for major copper ore types. Economic Geology 103, 615–628.

Hitachi, 2022. Ex5600-7b, t4f mining excavator. https://www.hitachicm.us/products/excavators/ex5600-7b/. Accessed: 2022-09-11.

Jara, R., Couble, A., Emery, X., Magri, E., Ortiz, J., 2006. Block size selection and its impact on open-pit design and mine planning. Journal of the Southern African Institute of Mining and Metallurgy 106, 205–211.

Journel, A., 1980. The lognormal approach to predicting local distributions of selective mining unit grades. Journal of the International Association for Mathematical Geology 12, 285–303.

Komatsu, 2019. Electric rope shovels p&h 4100xpc. https://mining.komatsu/es/product-details/p-h-4100xpc. Accessed: 2022-09-11.

Kurth, H., 2021. Digitalisation of ore composition and its benefits. https://www.ausimm.com/bulletin/bulletin-articles/digitalisation-of-ore-composition-and-its-benefits/. Accessed: 2022-09-11.

Lee, S.Y., Carle, S.F., Fogg, G.E., 2007. Geologic heterogeneity and a comparison of two geostatistical models: Sequential gaussian and transition probability-based geostatistical simulation. Advances in water resources 30, 1914–1932.

Li, G., Klein, B., He, C., Yan, Z., Sun, C., Kou, J., 2021. Development of a bulk ore sorting model for ore sortability assessment—part ii: Model validation and optimisation. Minerals Engineering 172, 107143.

Li, G., Klein, B., Sun, C., Kou, J., 2022. Investigation on influential factors of bulk ore sortability based on fractal modelling. Minerals Engineering 177, 107362.

Li, G., Klein, B., Sun, C., Kou, J., Yu, L., 2019. Development of a bulk ore sorting model for sortability assessment. Minerals Engineering 141, 105856.

Moss, A., Klein, B., Nadolski, S., 2018. Cave to mill: improving value of caving operations, in: Caving 2018: Proceedings of the Fourth International Symposium on Block and Sublevel Caving, Australian Centre for Geomechanics. pp. 119–132.

Nadolski, S., Klein, B., Elmo, D., Scoble, M., Liu, Y., Scholar, J., 2016. Investigation into the implementation of sensor-based ore sorting systems at a block caving operation. Proceedings of MassMin 2016 , 393–399.

Pizarro Munizaga, S.H.A., 2011. Modelamiento geoestadistico de leyes de cobre total y soluble. Master's thesis. Universidad de Chile.

Rendu, J.M., 2002. Geostatistical simulations for risk assessment and decision making: the mining industry perspective. International Journal of Surface Mining, Reclamation and Environment 16, 122–133.

Rossi, M.E., Deutsch, C.V., 2013. Mineral resource estimation. Springer Science & Business Media.

Sanhueza Passache, A.R., 2021. Evaluación económica de técnicas de selectividad de mineral (Ore Sorting) para su aplicación en Collahuasi. Master's thesis. Universidad de Chile.

Sturla-Zerene, G., Figueroa, E., Sturla, M., 2020. Reducing ghg global emissions from copper refining and sea shipping of chile's mining exports: A world win-win policy. Resources Policy 65, 101565.

Svanberg, A., Larsson, S., Mäki, R., Jonsén, P., 2021. Full-scale simulation and validation of bucket filling for a mining rope shovel by using a combined rigid fe-dem granular material model. Computational Particle Mechanics 8, 825–843.

Vann, J., Bertoli, O., Jackson, S., 2002. An overview of geostatistical simulation for quantifying risk, in: Proceedings of Geostatistical Association of Australasia Symposium" Quantifying Risk and Error, Citeseer. p. 1.

Vasylchuk, Y., Deutsch, C., 2018. Improved grade control in open pit mines. Mining Technology 127, 84–91.

de Werk, M., Ozdemir, B., Ragoub, B., Dunbrack, T., Kumral, M., 2017. Cost analysis of material handling systems in open pit mining: Case study on an iron ore prefeasibility study. The Engineering Economist 62, 369–386.

Yahyaei, M., Banisi, S., Hadizadeh, M., 2009. Modification of sag mill liner shape based on 3-d liner wear profile measurements. International Journal of Mineral Processing 91, 111–115.

Spatial multivariate morphing transformation on geochemical data augmentation[1]

Tong Li (tong.li@queensu.ca)
Julian M Ortiz (julian.ortiz@queensu.ca)

Abstract

Geochemical data plays an important role in supporting mineral deposits explorations in various ways. With the growth of deep learning methods utilized in mineral explorations, the demand for more geochemical data with more granularity keeps increasing. In contrast, the acquisition of geochemical data is usually expensive and time-consuming. In this work, we attempt to apply the spatial multivariate morphing transformation on geochemical data for data augmentation. This method decorrelates geochemical data both spatially and statistically by mapping the data into a multi-Gaussian space. Results show that this method is effective on geochemical data, but there are still problems that might be caused by the low stationarity and high dimensionality of geochemical data.

1. Introduction

Geochemical data is one of the most critical data that are used in mineral deposit exploration, especially in regional prospectivity mappings. Using geochemical data can provide vital information for discovering unknown mineral deposits like the spatial association of geochemical patterns, inter-elemental relationships, and geochemical anomalies that are caused by mineralization (Zuo and Xiong, 2020). Various methods had been adopted to process geochemical data to extract mineralization information. Among them, deep learning methods show a strong ability to intelligently find hidden patterns and features in geochemical data. The promise of such ability comes from large amounts of training data, and existing geochemical data usually cannot meet the demands and need augmentation. Traditional data augmentation methods designed on image data could result in severe problems when applied to geochemical data, like creating patterns in the wrong direction or adjusting all elements in the same scale.

To augment geochemical data, spatial multivariate morphing transformation (SMMT) (Avalos and Ortiz, 2022; Avalos et al., 2022) is utilized in this study. SMMT takes randomly sampled data from the initial dataset and statistically decorrelates them by mapping them from the initial multivariate space into a multi-Gaussian space. The decorrelated variables then are simulated with the spatial structure of random Gaussian values by geostatistical simulation algorithm. The simulated data are back-transformed by interpolating from the multi-Gaussian space back into the initial space. The interpolated data have a great reproduction of the multivariate features and relationships of the initial dataset and hence can be used as augmented data. SMMT provides a method to augment geochemical data that not only enlarges the variability of the initial data, but also honors the histogram and spatial variabilities of the initial data.

[1] Cite as: Li T, Ortiz JM (2022) Spatial multivariate morphing transformation on geochemical data augmentation, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-08, 119-131.

In this article, we summarize the application of SMMT in geochemical data augmentation, providing the workflow of SMMT in augmentation, we show results under different conditions, and discuss about problems we are facing and some implementation details.

2. A brief introduction to the geochemical data

The geochemical data used in this study is collected from the National Geochemical Surveying and Mapping Project of China (Wang et al., 2007; Xie et al., 1997). Standardized stream sediment samples from southern Jiangxi in China with an average sampling density of 1 sample/km^2 were taken and considered as the average geochemical concentration within this sample area (Xie, 1978). A total of 25 elements can be used in public studies, containing 7 major elements and 18 trace elements. All these elements are preprocessed by removing void values and negative values. All the data is located in a 335 × 335 grid, each grid cell represents 1km × 1km. Fig.1 shows the concentration of iron in the study area.

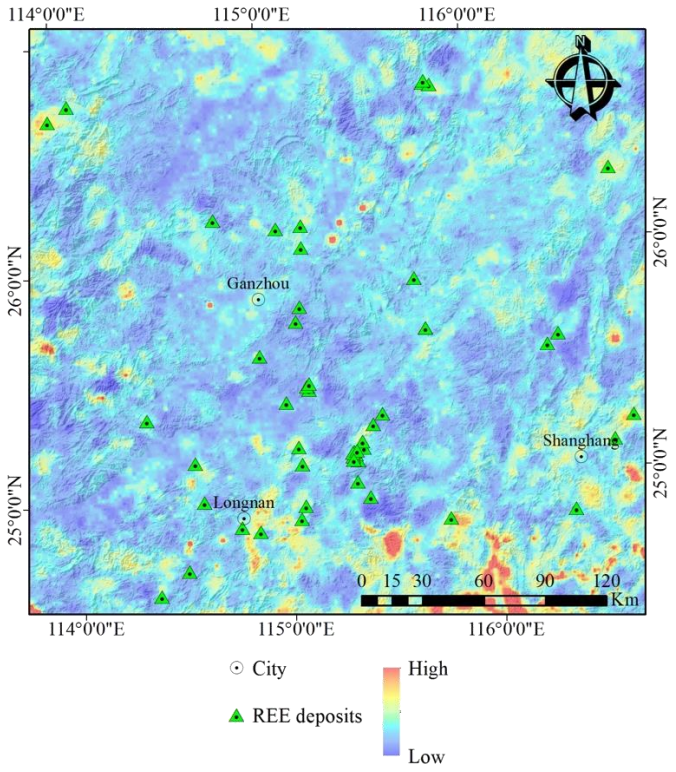


Figure 1: Geochemical distribution of iron (Li et al., 2022).

3. Spatial Multivariate Morphing Transform

In this section, we mainly focus on the workflow of applying SMMT to geochemical data, the theoretical background of SMMT can be reached in Avalos and Ortiz (2022).

1. **Choosing landmark points randomly from the original data.** Generate unduplicated locations from the grid of the study area as the landmark points for one iteration. Geochemical data on landmark points are normalized with an average of 0 and with a standard deviation of 1. Calculate the omnidirectional direct-variogram of each element and the cross-variogram between elements.

2. **Generating Morphing factors and pairing with landmark points.** Draw morphing factors for each element independently from a standard multi-Gaussian distribution with the same amount of landmark points. Compute the empirical cumulative distribution functions (CDF) of both landmark points and morphing factors. Pair the cumulative probability values of landmarks and the morphing factors with optimal transport. Optimal transport computes the Euclidean distances of every pair of samples on each dimension and attempts to find the optimal pair with the shortest distance over all dimensions. Calculate the omnidirectional variogram of paired morphing factors.

3. **Calculate the average of variograms of generated morphing factors.** Repeat Step 2 *n* times (100 times in this study). Calculate the average of all the omnidirectional variograms of morphing factors. The average variogram will be taken as the variogram model for sequential Gaussian simulation in Step 4.

4. **Sequential Gaussian simulation.** Simulate each set of generated morphing factors independently using the sequential Gaussian simulation (Goovaerts, 1997) for *m* times (100 times in this study). Compute the direct- and cross-variograms of the simulated data and make sure such variograms characterize the spatial structure of the morphing factors.

5. ***K*-nearest thin plate spline interpolation.** Map the simulated data from Gaussian space of value range (-∞, +∞) into logit space of value range (0, 1), to be conditioned to the cumulative probability values of landmark points of the value range of (0,1). For each simulated point, *k*-nearest landmark points are utilized as the control points in the thin plate spline interpolation (Bookstein, 1989). After interpolation of every point except for the landmark points, the result is considered as one augmentation of the original geochemical data. Repeat from step 1 for further augmentation until all the original points are used as landmark points.

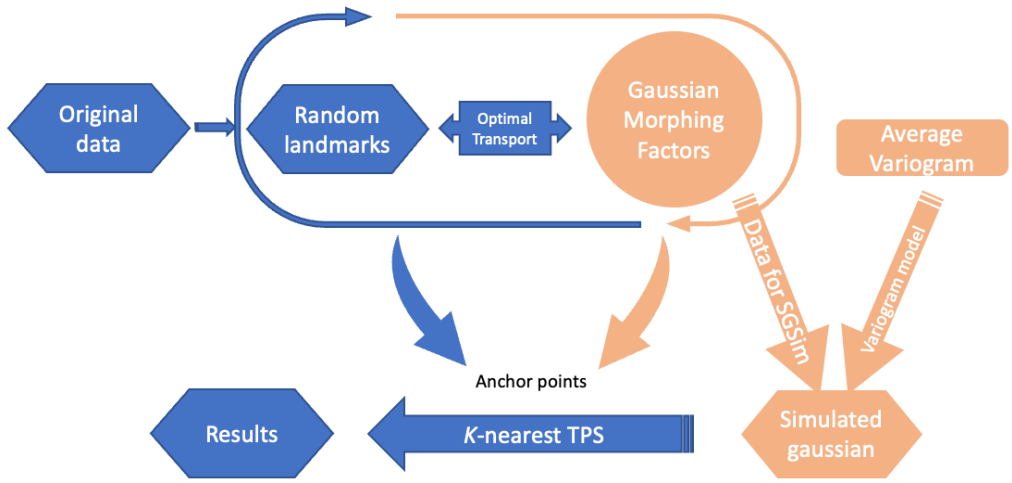


Figure 2: Schematic diagram of the workflow of the SMMT.

4. Results

The workflow in section 3 has been applied to three different sets of geochemical data for specific reasons. The subset with two dimensions contains two elements, iron and manganese, to intuitively illustrate the effectiveness of the SMMT; the subset of 25 dimensions contains the whole geochemical data, to test the SMMT on high dimensional data without extra preprocessing; and the subset with 17 dimensions is the production of preprocessing that removes the elements that are not correlated with other elements. Next, we show the results of the two-dimensional and 17-dimensional cases. Results from the 25-dimensional case are similar to those of the 17-dimensions.

4.1. Subset with 2 dimensions

Only 2 elements of the geochemical data are selected for illustration of the application of the SMMT on geochemical data. The omnidirectional direct-variograms and cross-variogram of two elements are shown in Fig.3. Spatial structures on both elements and between elements are observed.

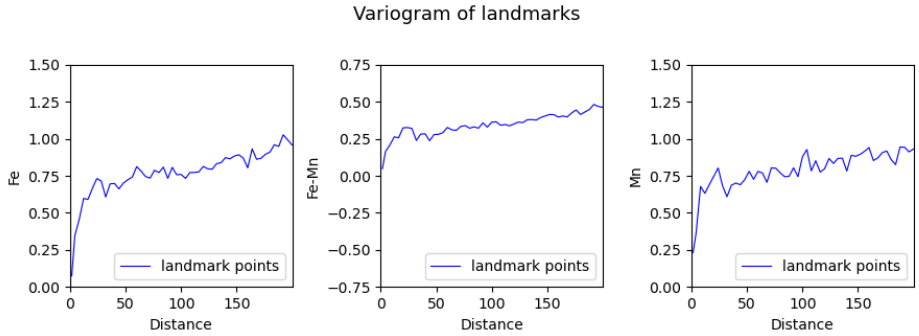


Figure 3: The omnidirectional direct-variograms and cross-variogram of two elements.

Fig.4 shows the landmark points (blue points) in the original space and the empirical cumulative distribution space and one set (out of 100 sets) of morphing factors (red points) in the Gaussian space and the empirical cumulative distribution space on the left side. The right side of Fig.4 shows part of the pairing of landmark points and morphing factors.

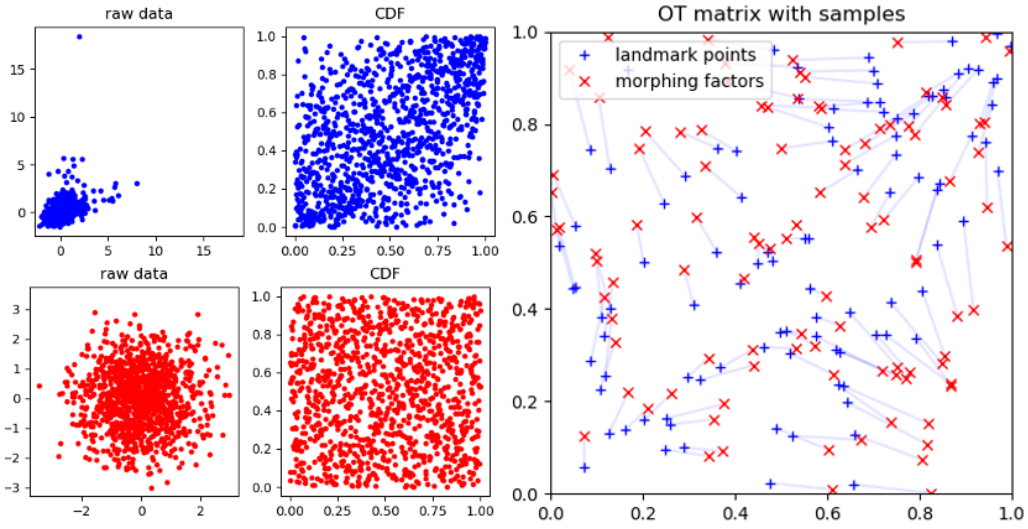


Figure 4: Pairing of the landmark points and morphing factors with 2-dimensional data.

After pairing with landmark points, the sequential Gaussian simulation is applied to every set of morphing factors with the average variogram model of the morphing factors. We use a maximum of 60 conditioning points with data assigned on nodes, and a search radius of 50 km. One realization of the simulation results is displayed in Fig. 5 (a). The spatial structure of the simulated data can be observed. The reproduction of variograms is shown in upper Fig.6. The direct variogram of iron and the cross variogram between iron and manganese are well-reproduced, but the direct variogram of manganese gets higher variance than the average variogram model.

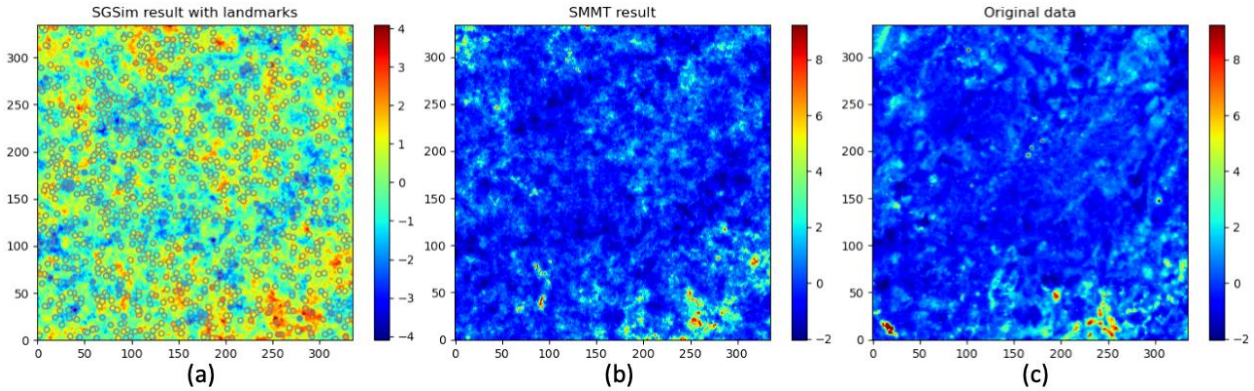


Figure 5: One realization of sequential Gaussian simulation, landmark points (dots in (a) with values), the respective result of SMMT and original data.

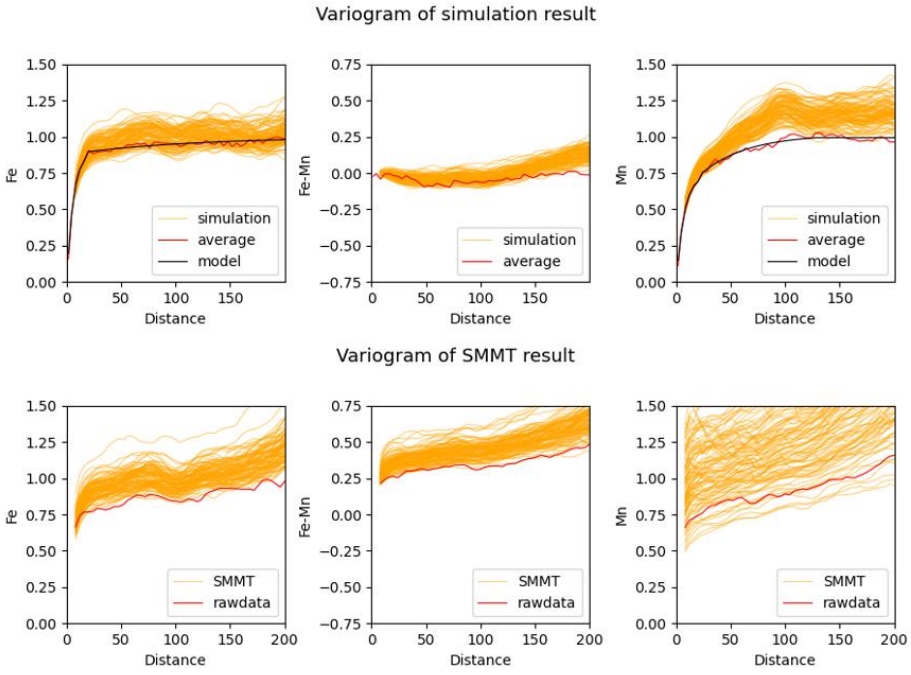


Figure 6: Variograms of the TPS results of 2-dimensional data.

The morphing factors are mapped into the original space via thin-plate spline interpolation (TPS). All points in the grid (except landmark points) are interpolated via TPS based on the 30 nearest landmark points. One of the results of the TPS is displayed in Fig.5 (b), which shows a similar spatial structure to the original data. The lower part of Fig. 6 shows the variogram of original geochemical data and the variogram

of the results of TPS. The variograms of TPS results show higher variances at most lag distances and the variograms of manganese are showing a similar structure but with unstable variances.

4.2. Preprocessed data with 17 dimensions

One of the purposes of utilizing the SMMT is to decorrelate the original geochemical data before simulation. However, high-dimensional data may disturb the optimal transport and affect the pairing. We reduce the dimension of geochemical data by removing some elements that are considered not correlated to other elements. 17 elements are left and used for augmentation with the SMMT. The omnidirectional direct-variograms and some of the cross-variograms are displayed in Fig.7 and Fig.8. All the elements are showing spatial structures with different variances.

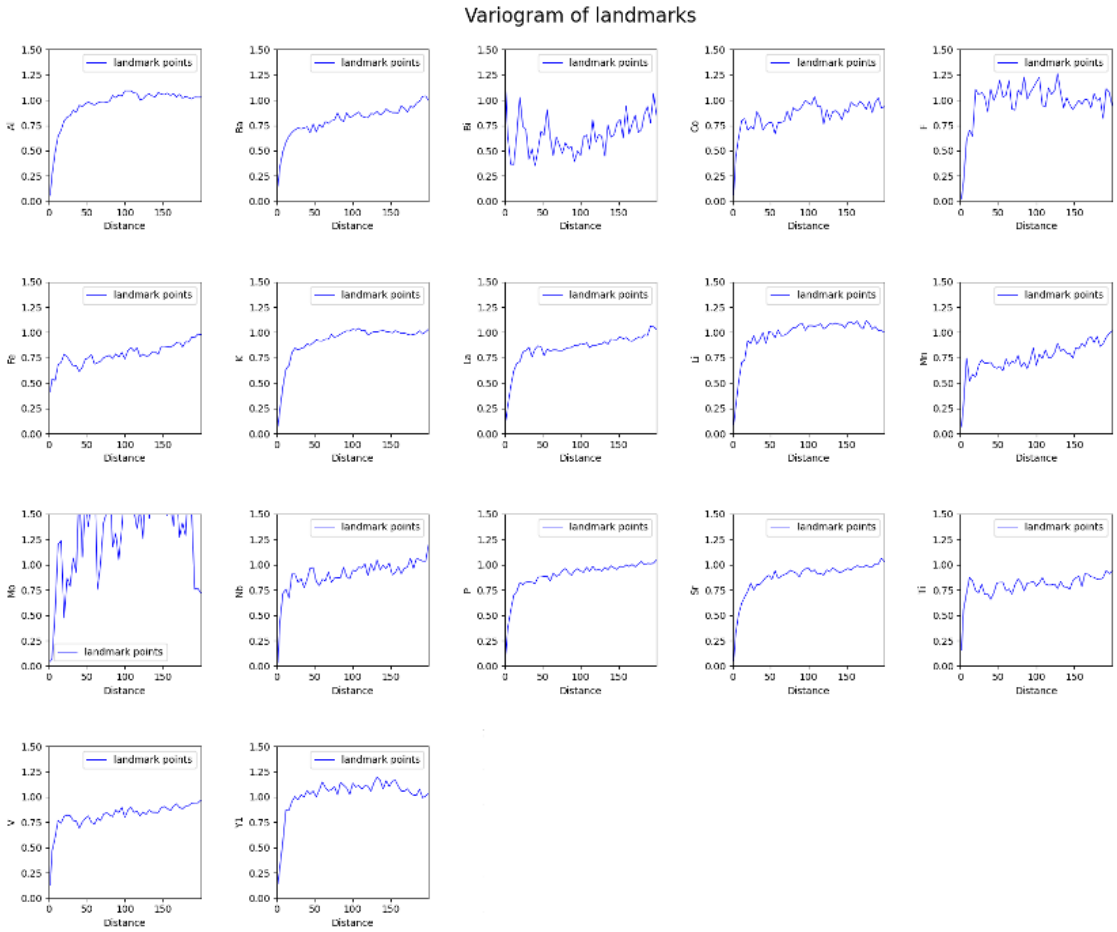


Figure 7: Omnidirectional direct variogram of landmark points.

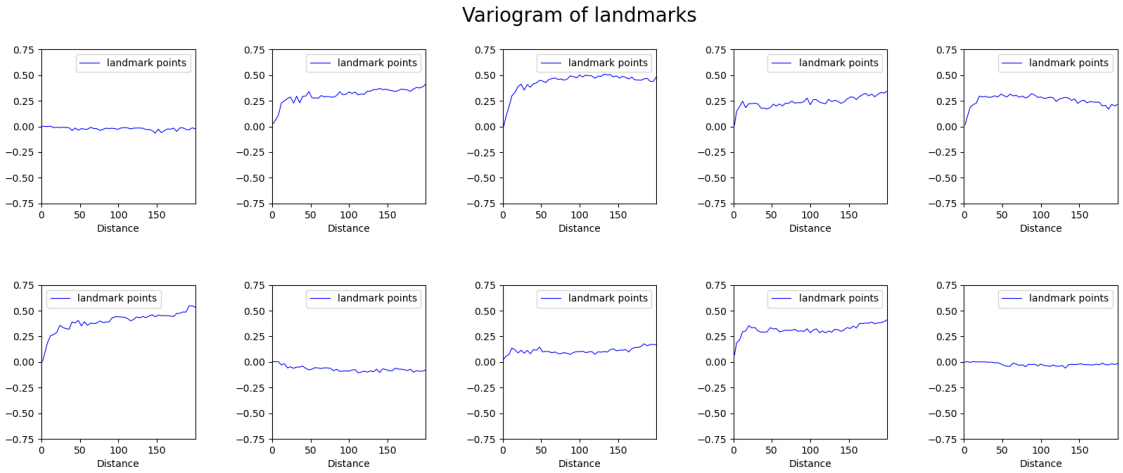


Figure 8: Omnidirectional cross variograms of landmark points (showing 10 out of a total of 136).

The pairing result of the iron and manganese of 17-dimension data is displayed in Fig. 9. With the higher dimension, the optimal transport of middle rankings is more cluttered than the top/last ranking in the cumulative probability values.

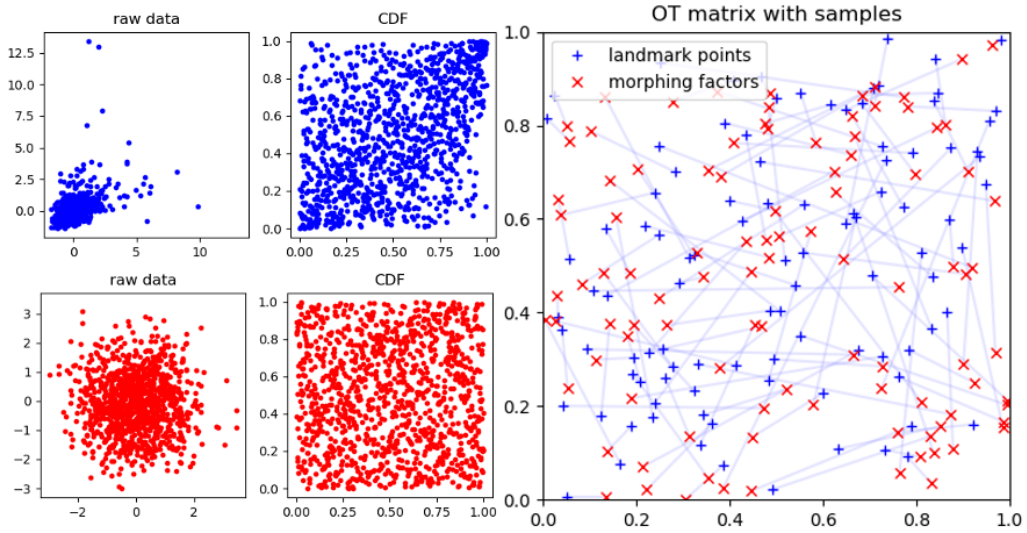


Figure 9: Pairing of the landmark points and morphing factors with 17-dimension data.

Fig. 10 (a) shows one realization of the sequential Gaussian simulation of the 17-dimension data. The simulation result exhibits poor spatial continuity with a fragmented spatial structure, which also resulted in high nugget effects in the direct-variograms shown in Fig. 11. The cross-variograms which are shown in Fig.12 indicate that all the elements are decorrelated.

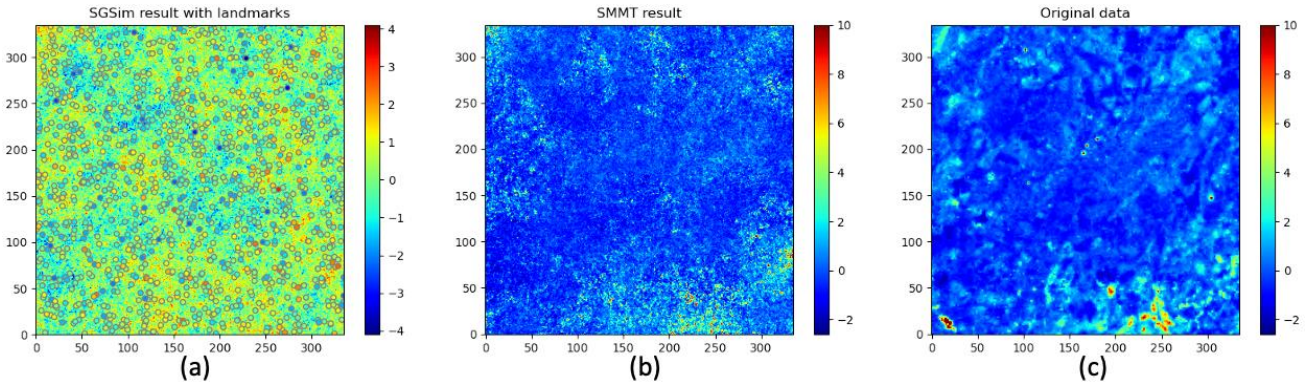


Figure 10: One realization of sequential Gaussian simulation, landmark points (dots in (a) with values), the respective result of SMMT, and original data.

Consequently, the TPS interpolations on such simulation results shown in Fig.10 (b), reproduce the trending of geochemical distribution on a large scale, but the spatial structure is not well-reproduced. Fig. 13 and Fig. 14 show the direct-variograms and cross-variograms of the results of the TPS interpolation. The spatial structure of the original data is reproduced but the average variances of results are mostly higher than those of the original data, which is similar to the 2-dimensional data.

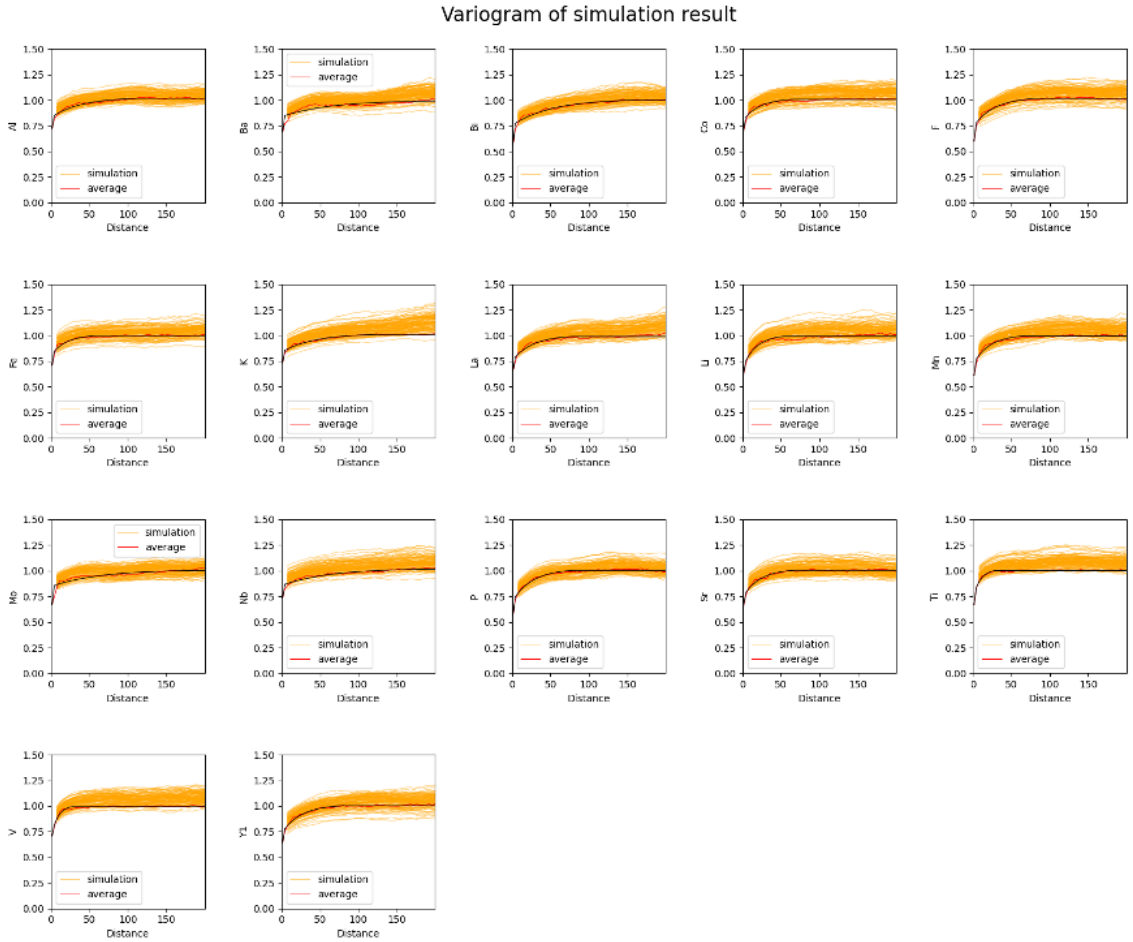


Figure 11: Direct-variograms of sequential Gaussian simulation results of 17-dimension data.

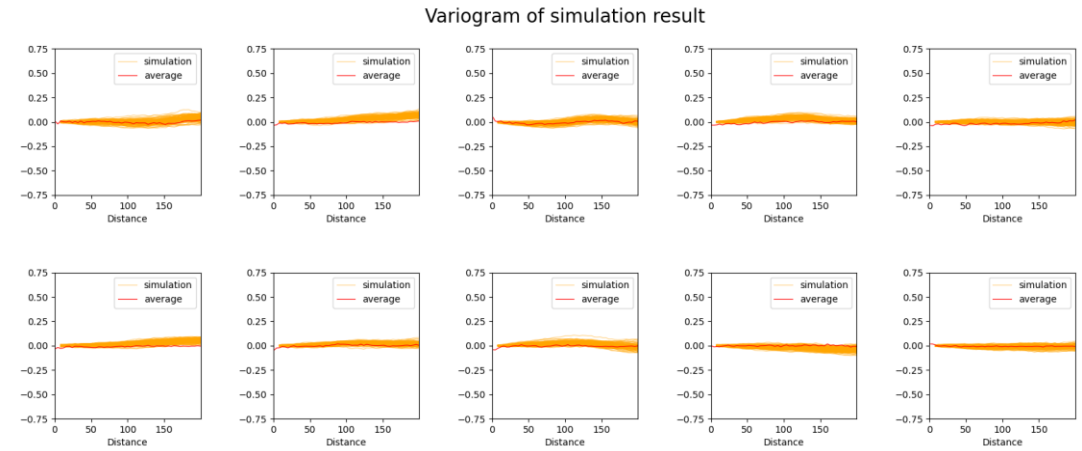


Figure 12: Cross-variograms of sequential Gaussian simulation results of 17-dimension data (showing 10 out of a total of 136).

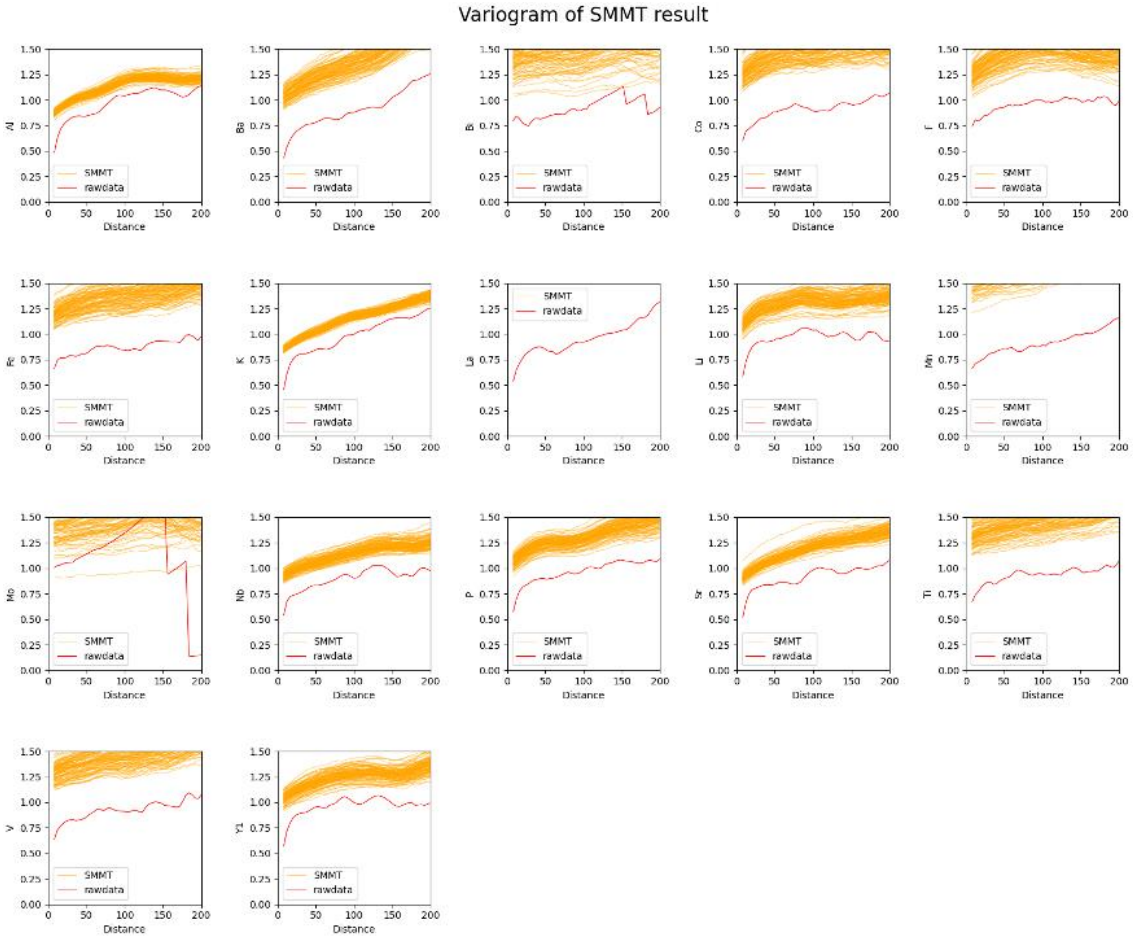


Figure 13: Direct-variograms of the TPS results of 17-dimension data.

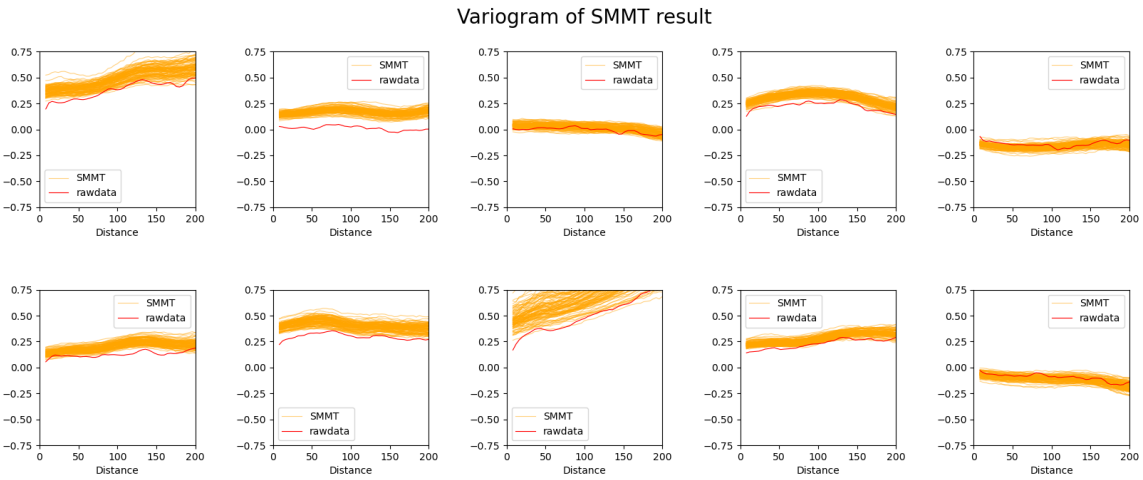


Figure 14: Cross-variograms of the TPS results of 17-dimension data (showing 10 out of a total of 136).

5. Discussion

Application of the SMMT on the different settings of geochemical data demonstrates that such a method can decorrelate the geochemical data and reproduce the spatial structures separately with a univariate geostatistical simulation algorithm. This section provides a discussion on several issues in applications on high-dimension data and some implementation details.

5.1. The curse of dimensionality

There are usually more than ten dimensions in geochemical datasets. Such high dimensionalities lead to a huge volume of data space and the available dataset is often inefficient to promise the effectiveness of algorithms.

In this study, the pairing in the optimal transport is sensitive to dimensionality. The optimal transport is trying to find the shortest distance to map data from the original space to a decorrelated Gaussian space while preserving their univariate relationships, which are their cumulative probabilities in this study. The shortest distance is the sum of distance on every dimension. When dimensions increase, the global optimal transportation on the whole dataset cannot preserve the univariate relationships on each dimension. Fig.15 shows the pairing results of a landmark point and its corresponding morphing factors with different dimensions. The ranking of the cumulative probability of landmark points becomes more unstable when the dimension increases.

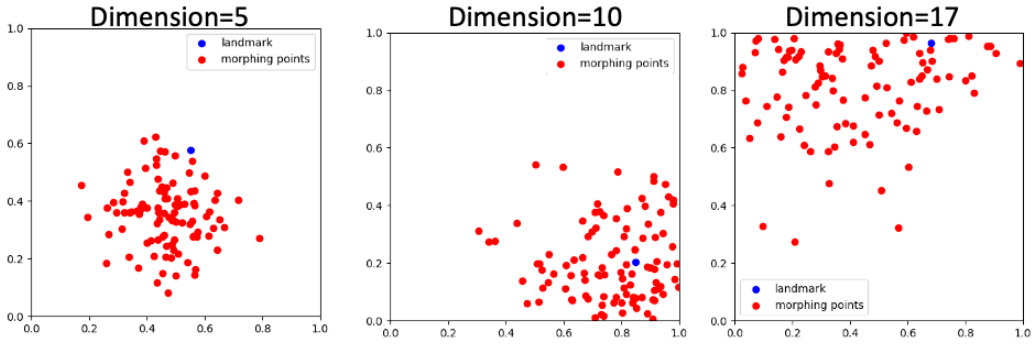


Figure 15: Pairing results of different dimensions.

The curse of dimensionality can be handled in two directions: The first is to increase the amount of available data, which is infeasible in this study; The other one is to restrict the data space by preprocessing the data to remove uncorrelated dimensions and various dimension reduction methods.

5.2. Landmark points selection

In geochemical data augmentation, the algorithm is expected to reproduce all the values and spatial structures of the original data. Therefore, all the points in the original data should be selected at least once as landmark points. The number of landmark points in the augmentation of geochemical data by the SMMT is a paradoxical parameter to consider. The more landmark points selected, the more values and spatial structures will be captured and reproduced, but the variability of augmented data becomes less.

5.3. Mapping simulated values

To project simulated Gaussian values back into the original data space, we use spatially *k*-nearest neighbors of the interpolated point as the anchor of projection in the TPS. This spatial interpolation of the simulated Gaussian values follows the 'First Law of Geography', which infers that near things are more related than distant things (Tobler, 1970). However, spatial interpolation may cause the high variances observed in Fig.6 and Fig.13 for not considering the statistical pattern of the landmark points. There is an alternative way as the statistical interpolation that projects the Gaussian values by their statistically *k*-nearest landmark points. The influences of statistical interpolation on geochemical data are worth exploring in future studies.

In the spatially *k*-nearest TPS interpolation, a different number of anchor points, *k*, produces a different result of interpolations (Fig.16). Theoretically, the number of anchor points must be larger than dimensions+1 while dealing with 2-D data (Rohr et al., 2001). With further increase of the anchor points, high values of landmark points show a stronger influence on neighborhoods. Therefore, the number of anchor points in TPS can be decided either based on the variogram of landmark points or domain knowledge of the geochemical data.

5.4. Implement details

In implementing the workflow of the SMMT, there are serval tricks to accelerate the procedures.

1. **Multi-processing of the sequential Gaussian simulation.** The sequential Gaussian simulation in GSLib can only simulate one dimension at one time. When simulating high-dimensional geochemical data, using the same executable program of the sequential Gaussian simulation with different parameter files will reduce the time on simulation.

2. **Finding *k*-nearest points in an adjustable window.** When projecting simulated Gaussian values back into the original data space, new interpolated points will be added into landmark points for better reproduction of the local structure. However, finding the k-nearest points of the simulated data requires the distances with all the other points which is a heavy calculation for the algorithm with a huge number of landmark points. A window of adjustable size is used to deplete such calculation. The size of the window is decided by the on-time density of the landmark points and to make sure there are more than *k* points inside.

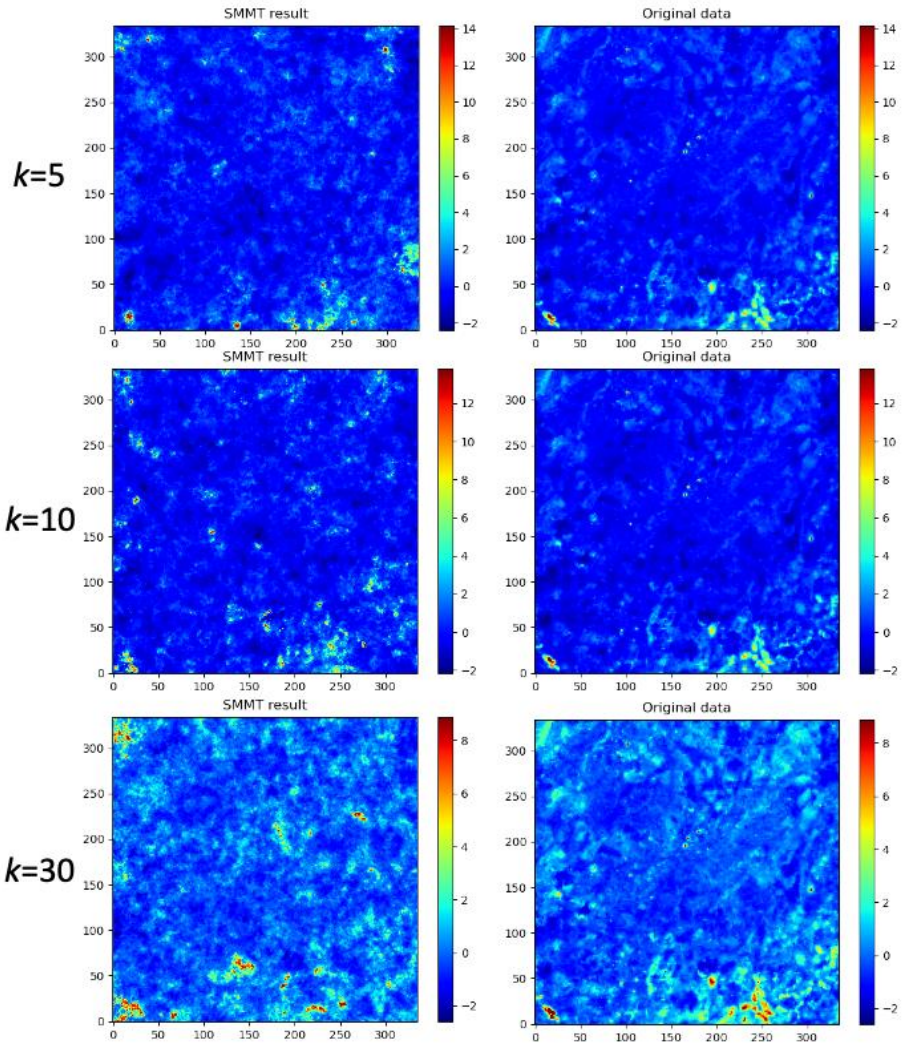


Figure 16: TPS results of different numbers of anchor points (k).

6. Conclusion

In this paper, we presented the workflow of the SMMT on geochemical data for data augmentation. Two sets of geochemical data with a different number of elements are utilized in the SMMT. Results of both sets of geochemical data have shown that the SMMT could decorrelate the data and reproduce the spatial structure of the geochemical data, but the variances of the augmented data were higher than the original data. Several problems that might cause this problem were discussed and future studies on improving the pairing in the optimal transport by reduction of data space and different ways of interpolation were issued to be done.

7. Acknowledgments

We acknowledge the support of the National Natural Science Foundation of China (No. 42172326).

8. References

Avalos S, Ortiz JM (2022) Spatial multivariate morphing transformation applied to geometallurgical attributes, in Geomet-Procemin 2022, Santiago, October 5-7 2022.

Avalos S, Ortiz JM, Leuangthong O (2022) Multivariate morphing transformation: Fundamentals and challenges, in 21st Annual Conference of the International Association for Mathematical Geosciences – IAMG 2022, Nancy, France, Aug 29-Sep 3, 2022.

Bookstein FL (1989) Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Transactions on pattern analysis and machine intelligence 11(6):567–585.

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press on Demand.

Li T, Zuo R, Zhao X, Zhao K (2022) Mapping prospectivity for regolith-hosted REE deposits via convolutional neural network with generative adversarial network augmented data. Ore Geology Reviews, 142, 104693.

Rohr K, Stiehl HS, Sprengel R, Buzug TM, Weese J, Kuhn M (2001) Landmark-based elastic registration using approximating thin-plate splines. IEEE Transactions on medical imaging 20(6):526–534.

Wang X, Zhang Q, Zhou G (2007) National‐scale geochemical mapping projects in China. Geostandards and Geoanalytical research, 31(4), 311-320.

Xie X (1978) Regional Geochemistry—National Reconnaissance Project. Bulletin of Geophysical and Geochemical Exploration, 3, 28.

Xie X, Mu X, Ren T (1997) Geochemical mapping in China. Journal of Geochemical Exploration, 60(1), 99-113.

Zuo R, Xiong Y (2020) Geodata science and geochemical mapping. Journal of Geochemical Exploration, 209, 106431.

Generative Adversarial Network 101[1]

Tong Li (tong.li@queensu.ca)
Julian M Ortiz (julian.ortiz@queensu.ca)

Abstract

Generative Adversarial Network, known as GAN, is one of the most promising architectures of generative models based on deep learning. The main purpose of the generative models is to generate synthesized data from the summarized distribution of given data. Therefore, finding the explicit or implicit approximation of the underlying real distribution of known data contributes the most to the ability of the generative models. With recent advances in deep learning algorithms, the impressive performance in recognizing the pattern of data further enhanced GAN's ability to synthesize data that fits in the distribution of real data. Such ability has great potential in the geoscience domain, in which the cost of new data is expensive. In this brief article, GAN, as a deep learning-based generative model within an adversarial framework, is introduced along with theoretical foundations, training procedures, and some variant structures of GAN.

1. Introduction

As we are going into the era of data, data science with deep learning algorithms are making striking achievements in lots of domains and changing our life. These data-driven methods have great abilities in discovering underlying hierarchical models that represent probability distributions over offered data (Bengio, 2009). The promise of such strong ability comes from a huge amount of input data training millions of parameters in these deep learning algorithms. In real applications like mining domains, data acquisition could be expensive and time-consuming. Researchers attempt to sidestep more data acquisitions by augmentations of known data, including translation, rotation, and flip (Krizhevsky et al., 2012). The diversity of data obtained by these minor modifications of data is relatively small. This motivates the utilization of generative models to produce synthetic data with more variability. Generative models approximate the real underlying probability distribution of input data and yield synthetic data within the best estimation of data distribution. However, difficulties in approximating many intractable probabilistic computations that caused by maximum likelihood estimation and related strategies hinder the performance of generative models (Goodfellow et al., 2014).

Generative Adversarial Network, as in its name, is a generative model trained in an adversarial net framework. GAN is normally composed of two networks, the discriminator and the generator. The discriminator is a classifier that tries to determine a sample more likely from real data distribution or GAN-modeled distribution and to separate synthetic data from real data. The generator learns to create synthetic data by incorporating feedback from the discriminator and tries to have a better estimation of the real data distribution to produce more realistic synthetic data. The basic principle of GAN is inspired by a two-player zero-sum game, in which each gain of utility is exactly the loss of the other player. The

[1] Cite as: Li T, Ortiz JM (2022) Generative Adversarial Network 101, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2022, paper 2022-09, 132-140.

competition between the discriminator and the generator promotes them to get trained and improve their methods until no one could ever win the game. When the competition terminates at an equilibrium point of the discriminator and the generator, which is called Nash equilibrium in game theory (Ratliff et al., 2013), GAN can be considered to have captured its best approximation of the distribution of offered data and hence can generate new data within that modeled distribution. Theoretically, any differentiable function that maps data from one space to another can be used as the discriminator and the generator. GAN is initially implemented by multilayer perceptrons and commonly by variants of neural networks nowadays.

In the following sections, we will introduce some theoretical foundations of GAN, GAN implementations of these foundations, the structure and training procedures of GAN, and we will end with a brief review of variants of GAN.

2. Theoretical Foundation of GAN

2.1. Estimation on distribution

Generative models aim to discover the underlying distribution $P_{data}(x)$ of offered data (also called probability density function) which is composed by samples $\{x^1, x^2, \dots, x^n\}$. In real practice, $P_{data}(x)$ is usually unknown and extremely intricate. The goal of generative models is to find the best approximation of $P_{data}(x)$, which can be defined by a set of parameters θ, denoted as $P_G(x;\theta)$. By the way of finding the best approximation of $P_{data}(x)$, generative models can be classified into two classes: explicit density model and implicit density model. An explicit density model assumes the distribution and utilizes true data to train the model containing the distribution or fit the distribution parameters. An implicit density model produces synthetic data without an explicit distribution and uses produced data to modify the model.

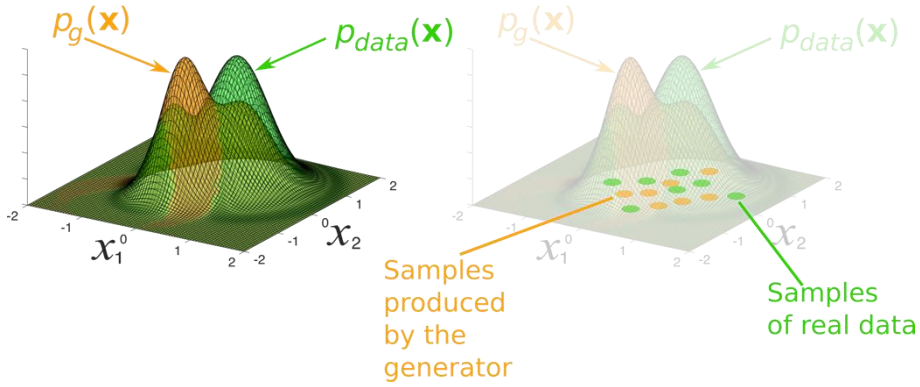


Figure 1: Illustration of estimation (orange) of real probabilities distribution (green) based on samples (Dumoulin et al., 2017).

To figure out how generative models find the best approximation of P_{data} utilizing a limited amount of real data, maximum likelihood estimation, as a representative of generative models, is illustrated. The likelihood function (Eq.1) is defined by the joint probability between real samples and the approximation distribution, which is used for the estimation.

$$L = \prod_{i=1}^{n} P_G(x^i;\theta) \tag{1}$$

To maximize the likelihood function, the best set of parameters θ^* is calculated as:

$$\theta^* = \arg\max_{\theta} L \tag{2}$$

The natural logarithm of the likelihood function is often used for calculation convenience. Since the logarithm is monotonic, the maximum of both functions occurs at the same value of θ.

$$\begin{aligned}\theta^* &= \arg\max_{\theta} \ln \prod_{i=1}^{n} P_G(x^i;\theta) \\ &= \arg\max_{\theta} \sum_{i=1}^{n} ln\, P_G\left(x^i;\theta\right) \\ &\approx \arg\max_{\theta} E_{x\sim P_{data}}[\log P_G\,(x;\theta)]\end{aligned} \tag{3}$$

Eq. 3 is the main indicator for training both the generator and the discriminator in GAN, as well as where the adversary happens, which is going to be introduced in the next section.

If we go further on Eq. 3, minus it with the probability distribution of real samples in their own distribution, which is independent of θ, as

$$\begin{aligned}\theta^* &\approx \arg\max_{\theta} E_{x\sim P_{data}}[\log P_G\,(x;\theta)] \\ &= \arg\max_{\theta} E_{x\sim P_{data}}[\log P_G\,(x;\theta)] - E_{x\sim P_{data}}[\log P_{data}\,(x)] \\ &= \arg\max_{\theta} \int_x P_{data}\,(x)\log P_G(x;\theta)dx - \int_x P_{data}(x)\log P_{data}(x)dx \\ &= \arg\min_{\theta} KL(P_{data}(x)\|\,P_G(x;\theta))\end{aligned} \tag{4}$$

in which KL denotes Kullback–Leibler divergence. The KL divergence measures how two probability distributions are different from each other (Kullback and Leibler, 1951). By minimizing the KL divergence, the generative model will reach the set of parameters of the best approximation of underlying data distribution.

2.2. GAN-implementations of estimation

Estimation of the data distribution is done by fitting a known parametrized distribution to an underlying intractable distribution. One of the most powerful tools in fitting nonlinear functions, a neural network, is used in GAN to implement such estimation. The generator and the discriminator in GAN adopt two independent neural networks with adversarial purposes:

The generator neural network, denoted as ***G***, is defined by θ in Eq.3. ***G*** takes random input z from a predefined simple distribution like uniform and gaussian distribution $P_z(z)$ as input and represents a mapping to data space $G(z;\theta_G)$. The purpose of ***G*** is to yield synthesized data that cannot be detected by the discriminator.

The discriminator neural network, denoted as ***D***, is a discriminative model that outputs a single scalar $D(x;\theta_D)$ trying to separate synthesized data from real data. The best set of parameters and purpose of ***D*** can be defined as

$$\theta_D{}^* = \arg\max_{\theta_D}(\, E_{x\sim P_{data}}[\log D\,(x;\theta_D)] + E_{z\sim P_z}[\log(1 - D\,(G(z;\theta_G);\theta_D))]) \tag{5}$$

As a summary, G and D play the following minimax game with value function $V(G,D)$:

$$\min_G \max_D V(G,D) = E_{x\sim P_{data}}[\log D\,(x;\theta_D)] + E_{z\sim P_z}[\log(1 - D\,(G(z;\theta_G);\theta_D))] \tag{6}$$

Given any generator G, the value function $V(G,D)$ can be reformulated as below:

$$\begin{aligned} V(G,D) &= \int_x P_{data}\,(x)\log D\,(x;\theta_D)dx - \int_z P_z(z)\log(1 - D\,(G(z;\theta_G);\theta_D))dz \\ &= \int_x P_{data}\,(x)\log D\,(x;\theta_D) + P_g(x)\log(1 - D\,(x;\theta_D))dx \end{aligned} \tag{7}$$

For any $(a,b) \in \mathbb{R}^2\,\{0,0\}$, the function $a\log(y) + b\log(1-y)$ achieves its maximum in [0,1] at $\frac{a}{a+b}$. Thus, the optimal discriminator ***D*** is given by the maximum of $V(G,D)$.

$$D^*{}_G = \frac{P_{data}(x)}{P_{data}(x) + P_g(x)} \tag{8}$$

With Eq. 8 and the KL divergence, the value function with a fixed ***G*** Eq.7 could be written as:

$$\begin{aligned} V(G,D) &= E_{x\sim P_{data}}[\log D^*{}_G\,(x)] + E_{z\sim P_g}[\log(1 - D^*{}_G\,(x))] \\ &= E_{x\sim P_{data}}\left[\log\frac{P_{data}(x)}{P_{data}(x) + P_g(x)}\right] + E_{z\sim P_g}\left[\log\frac{P_d(x)}{P_{data}(x) + P_g(x)}\right] \\ &= KL\big[(P_{data}\|(P_{data} + P_g)\big] + KL\big[\big(P_g\|(P_{data} + P_g)\big] \end{aligned} \tag{9}$$

Therefore, the value function of GAN is proven related to the KL divergence that measures the likelihood of two probabilistic distributions. Competition in this minimax game leads both the generator and discriminator to keep optimizing their methods. Finally, based on Sion's minimax theorem (Sion, 1958) and the proof in Goodfellow et al. (2014), two players ***G*** and ***D*** will converge at the same point, where $P_g = P_{data}$, then $G(z;\theta_G{}^*)$ can be considered as the best estimation of P_{data} and can be utilized to generate synthetic data.

3. Common structure and training of GAN

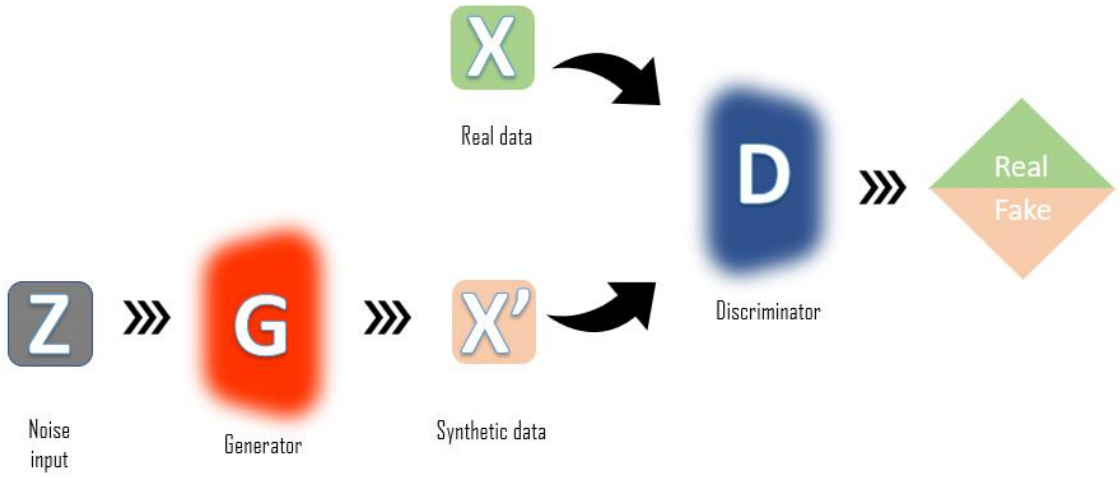


Figure 2: Typical GAN structure.

GAN is usually comprised of two core models, the generator and the discriminator. The generator takes random inputs and maps them into synthetic data with a differentiable function that can be trained by backpropagation. These random inputs are usually generated from a simple distribution like Gaussian distribution but can also be generated by certain rules in some variants of GAN for better estimation of the real data distribution or other purposes. The discriminator takes samples either from real data or from the generator and predicts a binary class label of real or synthetic. The results derived by the discriminator will be evaluated by ground truth and used for guiding optimization of the weaker model in the game.

Detailed training procedures are described as:

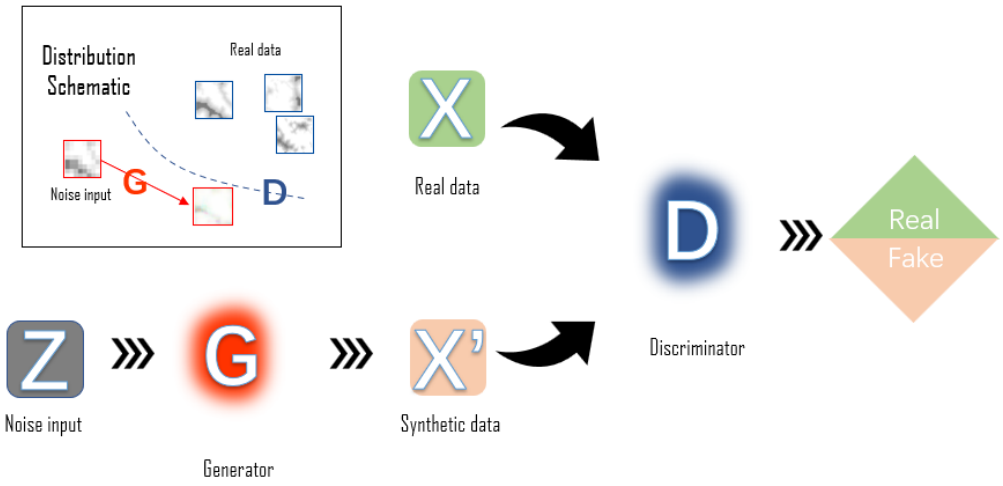


Figure 3a: The beginning of training a GAN.

Step one: At the very beginning, the generator has been set up randomly and hence the synthetic can easily be distinguished by the discriminator. The optimizer of GAN will train the generator by the gradients of the value function (or loss function in machine learning).

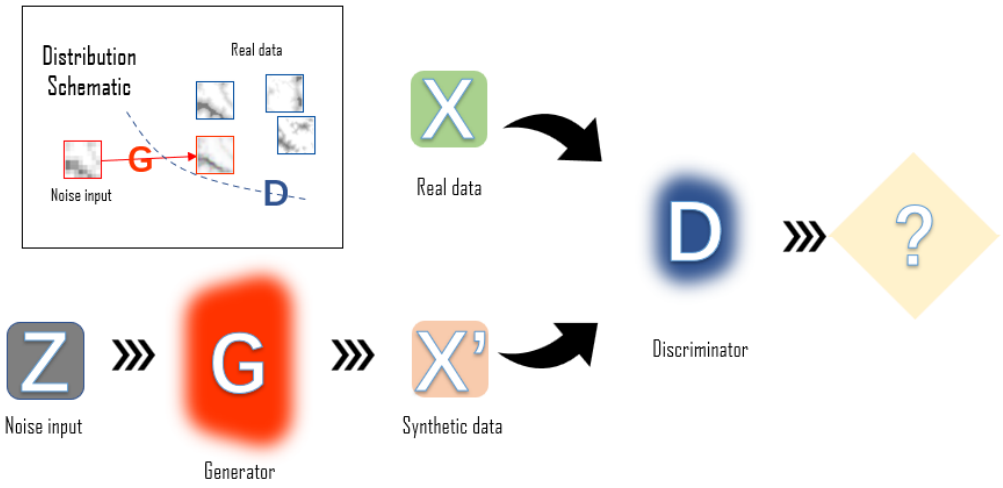


Figure 3b: After training the generator, the discriminator cannot decide which is fake.

Step two: Once the generator gets trained, the discriminator should be trained at least once. The reason for this procedure is a more optimized discriminator can provide a more accurate ratio between 2 distributions (see Eq. 8) and hence produce better gradients to train the generator.

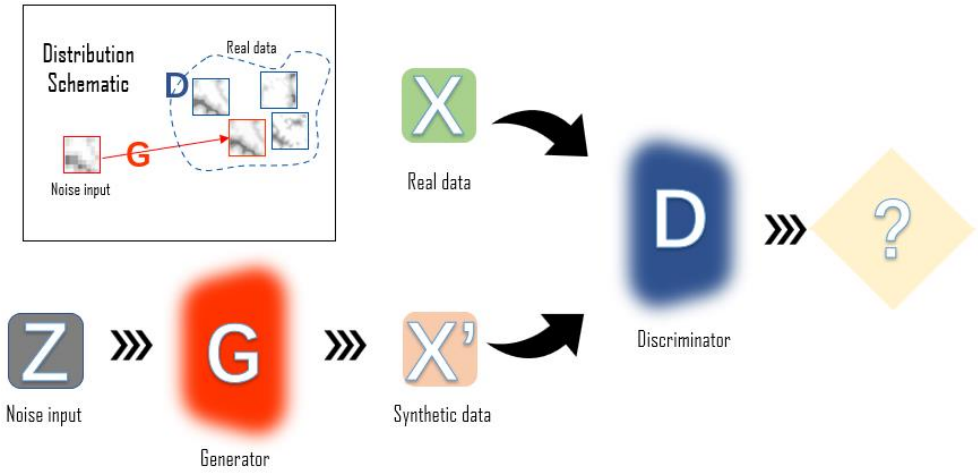


Figure 3c: In the final stage of the training of GAN, both the generator and the discriminator cannot be further trained.

Step three: Keep training with steps one and two, until the discriminator is maximumly confused and can not distinguish synthetic data from real data, which reflects in the prediction of 0.5 of any input sample either real or synthetic. Then the training is considered finished, and the synthetic data produced by the generator is lying in the distribution of real data.

The training procedure for GAN is usually challenging and unstable due to several reasons (see Radford et al.,2015 and Salimans et al., 2016). Certain useful tricks and improvements in the training of GAN are used:

1. Batch Normalization. Using a large batch of real or synthetic data after normalization for both the discriminator and generator will stabilize the training and speed up training.

2. High learning rate and more steps on training the discriminator. As mentioned in the training procedure, focusing more on the discriminator will provide with better gradient for the whole network.

3. Avoid hard boundaries. Deep neural nets are prone to producing highly confident outputs that identify the correct class but with too extreme of a probability (Goodfellow et al., 2016). Smoothing the boundaries of the activation functions by using leaky ReLU in the neural network (see Radford et al., 2015 for details) and using a one-side smoothed label (usually setting 0.9 instead of 1 on positive samples) would help the discriminator more efficiently on distinguishing of synthetic data (Salimans et al., 2016).

4. Variants of GAN

Since the original structure of GAN proposed by Goodfellow et al. in 2014, a huge number of GAN variants have been created. These variants of GAN are innovative for their improvement of structures, extensions of theory, or specific application-oriented design.

4.1. Wasserstein GAN

Instead of using KL divergence as the value function for optimization, Arjovsky et al. (2017) proposed Wasserstein GAN that uses the Earth-Mover distance for evaluating the distribution distance between real data and the synthetic data and uses a critic function supported by Lipschitz constraint to represent the discriminator. The advantage of using such Wasserstein distance is to stabilize the training of GAN by avoiding the gradient vanishing problem when the real data and synthetic data share very little overlap.

4.2. Deep convolutional GAN

The original structure of GAN uses multilayer perceptron (MLP) as the generator and discriminator. The limitation of MLP is all the data is treated the same and does not leverage the benefit of spatial structures as the convolutional neural network (CNN) does. Thus, CNN has way better performance at generating images than MLP. Deep convolutional GAN (DCGAN, Radford et al., 2016) substitutes the CNN for MLP in the original GAN and soon became the backbone of most of the variants of GAN. CNN structures in DCGAN utilize convolutional layers to capture features of data on different scales with convolutional kernels, the generator produces synthetic data with these features while the discriminator captures its own set of features that can be used to distinguish synthetic data from real ones.

4.3. Conditional GAN and Info GAN

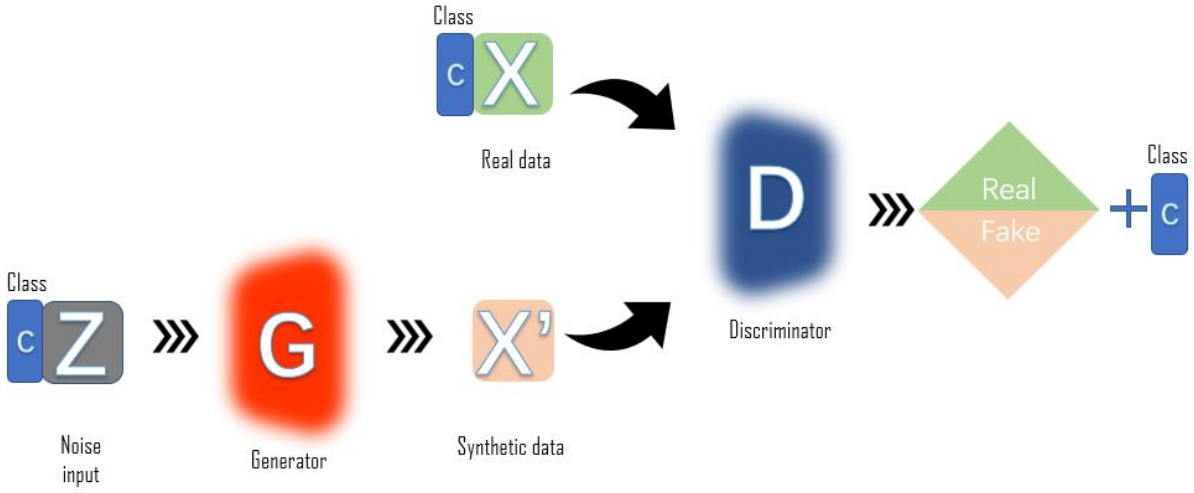


Figure 4: Intuitive structure of Info GAN.

The original GAN can only discover one data distribution and generate data in one estimated distribution. Conditional GAN (Mirza and Osindero, 2014) and Info GAN (Chen et al., 2016) add a "latent code" to the random noise input of GAN and adapt the generator and discriminator to this latent code in outputting class label of synthetic data. The value function of Info GAN is formulated as below:

$$\min_G \max_D \{f_I(D,G) = f(D,G) - \lambda I(c; G(z,c))\} \quad (10)$$

4.4. Super-resolution GAN

Super-resolution GAN (Ledig et al., 2017) takes low-resolution images as input instead of random noise and generates realistic details of the images while up-sampling. To achieve photorealistic super-resolution, SRGAN utilizes a perceptual loss function (notes as value function above) which consists of an adversarial loss to make sure the image manifold and a content loss to assure the perceptual similarity between real data and synthetic data other than simply pixel similarity. SRGAN uses ResNet (see He et al., 2016 for

details) in the generator, which is noted for its residual block that adds former output with its own output, called skip connection. ResNet in SRGAN is easier than training a CNN for capturing features and allows SRGAN to be substantially deeper to generate better results.

5. Conclusions

In this paper, we elaborate on the theoretical foundations, algorithm implementations, training procedures with useful tricks, and common variants of GAN. GAN is an implicit generative model that utilizes two neural networks in competition to generate synthetic data fitted in the underlying distribution of real data. The ability of GAN to generate infinite synthetic data from limited real data has great application values and potential in the geoscience domain.

6. Acknowledgments

We acknowledge the support of the National Natural Science Foundation of China (No. 42172326).

7. References

Arjovsky M, Chintala S, Bottou L (2017, July) Wasserstein generative adversarial networks. In *International conference on machine learning* (pp. 214-223). PMLR.

Bengio Y (2009) Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, *2*(1), 1-127.

Chen X, Duan Y, Houthooft R, Schulman J, Sutskever I, Abbeel P (2016) Infogan: Interpretable representation learning by information maximizing generative adversarial nets. *Advances in neural information processing systems*, *29*.

Goodfellow I (2016) Nips 2016 tutorial: Generative adversarial networks. *arXiv preprint arXiv:1701.00160*.

Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Bengio Y (2020) Generative adversarial networks. *Communications of the ACM*, *63*(11), 139-144.

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90.

Kullback S, Leibler RA (1951) On information and sufficiency. *The annals of mathematical statistics*, *22*(1), 79-86.

Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Shi W (2017) Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681-4690).

Mirza M, Osindero S (2014) Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

Ratliff LJ, Burden SA, Sastry SS (2013) Characterization and computation of local nash equilibria in continuous games. In Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on, pages 917–924. IEEE.

Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X (2016) Improved techniques for training gans. In Advances in Neural Information Processing Systems, pages 2226–2234.

Sion M (1958) On general minimax theorems. *Pacific Journal of mathematics*, *8*(1), 171-176.