Artificial Intelligence Algorithms for Real-time Production Planning with Incoming New Information in Mining Complexes

Ashish Kumar

Doctor of Philosophy

Mining and Materials Engineering

McGill University
Montreal, Quebec
2020-07-06

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

ACKNOWLEDGEMENTS

Finally, I would like to thank my family for their continuous support during the Ph.D. To my dad, for his motivation, unwavering support, and for his presence – you could not be there in person, but you were never more than a phone call away. To my sister, Ankita, and my brother, Abhishek, for keeping my spirits up with their stories and constant assurances that all would be well. A special thanks to my partner, Karine, for being there through the ups and downs and for staying level-headed when I could not be, and for her help in editing and translating parts of this thesis.

This may be *my* Ph.D., but it has been *our* journey.

CONTRIBUTION OF AUTHORS

The author of this thesis is the primary author in all the papers included in the related chapters. Professor Roussos Dimitrakopoulos was the supervisor and is included as co-author in all four papers of this thesis. Marco Maulen is the third author of the second paper and provided comments on the application of the method proposed in that paper. The four papers are:

- "Application of simultaneous stochastic optimization with geometallurgical decisions at a copper-gold mining complex. Kumar, A., and Dimitrakopoulos, R. (2019), Mining Technology, 128(2), pp. 88-105.

- "Adaptive self-learning mechanisms for updating short-term production decisions in an industrial mining complex." Kumar, A., Dimitrakopoulos, R., and Maulen, M. (2020), Journal of Intelligent Manufacturing, pp. 1-17, doi: https://doi.org/10.1007/s10845-020-01562-5.

- "Production scheduling in industrial mining complexes with incoming new information using tree search and deep reinforcement learning." Kumar, A., and Dimitrakopoulos, R. (2020), Submitted to Applied Soft Computing.

- "Updating geostatistically simulated models of mineral deposits in real-time with incoming new information using actor-critic reinforcement learning." Kumar, A., and Dimitrakopoulos, R. (2020), Submitted to Computers and Geosciences.

ABSTRACT

A mining complex is an integrated mineral value chain that consists of multiple interrelated activities, from material extraction to a set of sellable products. Simultaneous stochastic optimization of the long-term plan of a mining complex provides asset valuation, strategic decisions, forecasts, and targets while accounting for geological supply uncertainty, and commodity price uncertainty. However, supply uncertainty related to non-additive geometallurgical properties of materials that affect the performance of the processing streams is still not accounted for in the existing simultaneous stochastic optimization models. A short-term plan developed within the long-term plan provides decisions on the monthly/weekly/daily sequence of extraction, the destination of materials, the allocation of equipment, and the utilization of processing streams, and aims to ensure compliance with the long-term targets while maximizing cash flow. With digital technologies, a mining complex can acquire new information about the quality and quantity of materials at, and performance of its different components while it operates. However, the new information is partial and noisy, therefore uncertain, and referred to as "soft" data. This "softness" is attributed to the characteristics of the related sensors that generate indirect measurements compared to the geochemical analysis of drillhole samples. Existing technologies cannot integrate the soft incoming new information and respond accordingly in real-time to adapt the short-term plan.

This thesis contributes a new self-learning framework for adapting short-term planning decisions by learning to integrate soft incoming new information. This follows an extension of the simultaneous stochastic optimization model to account for the geometallurgical supply uncertainty.

This thesis first introduces an approach that characterizes the materials as hard or soft, then links them to the simultaneous stochastic optimization model through geometallurgical targets. An application at a copper-gold mining complex indicates higher chances of meeting targets, a substantial increase in metal production, and a 19.3% increase in net present value.

Next, a self-learning framework is presented that proposes a new extension of the ensemble Kalman filter for updating the supply uncertainty of multiple correlated properties of materials with incoming new information. Then, a model-free policy gradient reinforcement learning algorithm is extended to learn to adapt the destination of materials in a multi-product mining complex. The proposed framework is applied at a copper mining complex and compared to industry-standard approaches to highlight its abilities.

A new model-based self-play reinforcement learning algorithm is proposed next, which uses a Monte Carlo tree search and a deep neural network agent in a self-play architecture to learn to respond to incoming new information by adapting the short-term sequence of extraction, destination of materials, and utilization of processing streams simultaneously. The algorithm plays the game of short-term production planning by itself using a Monte Carlo tree search to train a deep neural network agent. This work also proposes a Monte Carlo simulation algorithm to update the equipment performance uncertainty. An application at a copper mining complex shows its ability to adapt the short-term plan almost in real-time to better meet production requirements while increasing metal production and cash flows.

The existing methods to update the supply uncertainty of materials cannot learn from the incoming new information and do not account for high-order spatial statistics. Therefore, a new model-free actor-critic reinforcement learning algorithm is proposed. First, the actor and critic agents respectively predict and evaluate the updated properties of a mining block, then the updated block properties are evaluated with a conditional probability distribution function

generated using high-order spatial statistics and, finally, the evaluations are used to train the agents. An application at a copper mining operation demonstrates its applied aspects.

The self-learning artificial intelligence framework proposed in this thesis can allow a mining complex to learn and adapt with incoming new information while it operates to make more informed short-term planning decisions. Future work can improve the performance of the algorithms and develop new methods to adapt the long-and short-term plans simultaneously.

RÉSUMÉ

Un complexe minier est une chaîne d'approvisionnement intégrée composée de plusieurs activités interreliées, allant de l'extraction de minerai, à la livraison de produits commercialisables. L'optimisation stochastique simultanée du plan de production à long terme d'un complexe minier fournit des décisions stratégiques, des prévisions, des objectifs, ainsi qu'une évaluation des actifs, tout en tenant compte des incertitudes géologiques, et de l'incertitude des prix des marchandises. Cependant, l'incertitude liée aux attributs géométallurgiques non-additifs de matériaux qui affectent la performance des réseaux de traitement n'est toujours pas pris en compte dans les modèles existants d'optimisation stochastique simultanée. Un plan de production à court-terme, créé pour fournir des décisions sur une base mensuelle/hebdomadaire/journalière dans le contexte d'un plan de production à long-terme, vise à assurer la conformité avec les objectifs à long terme, tout en maximisant la liquidité. Avec des technologies numériques, un complexe minier peut acquérir de nouvelles informations pendant son opération concernant la performance de ses différentes composantes, ainsi que la qualité et la quantité de matériaux qui s'y retrouvent. Cette information qui, toutefois, est partielle et bruitée, donc incertaine, est connue sous le nom de données « souples ». La « souplesse » de cette nouvelle information est due aux caractéristiques des capteurs associés qui génèrent des mesures indirectes, plutôt que celles provenant de l'analyse géochimique d'échantillons de forage. Les technologies existantes ne peuvent pas incorporer l'information entrante bruitée et incertaine, et y répondre adéquatement en temps réel pour adapter la planification à court-terme.

La contribution de cette thèse est de proposer un nouveau cadre d'auto-apprentissage qui permet d'adapter les décisions de planification à court terme en apprenant à intégrer

l'information entrante bruitée et incertaine. Ceci s'ajoute à une extension du modèle d'optimisation stochastique simultanée pour prendre en compte l'incertitude géométallurgique des ressources.

Cette thèse introduit d'abord une approche qui caractérise les matériaux comme durs ou mous, puis les relie au modèle d'optimisation stochastique simultanée à travers des objectifs géométallurgiques. Une application dans un complexe minier de cuivre-or démontre de plus grandes chances d'atteindre les différents objectifs de production, une augmentation considérable de production de métaux, et une hausse de 19.3% de la valeur présente nette.

Ensuite, un cadre d'auto-apprentissage est présenté, qui propose une extension du filtre de Kalman d'ensemble pour mettre à jour l'incertitude d'approvisionnement d'une multitude d'attributs corrélés de matériaux avec la nouvelle information entrante. Par la suite, un algorithme de gradient de politique d'apprentissage par renforcement sans modèle est appliqué pour apprendre à répondre à la nouvelle information entrante dans un complexe minier à multiples produits. Le cadre proposé est appliqué dans un complexe minier de cuivre, et comparé aux approches standards de l'industrie afin de démontrer son habileté.

Un nouvel algorithme d'apprentissage par renforcement auto-jouant basé sur un modèle est ensuite proposé, qui utilise une recherche arborescente de Monte Carlo et un agent de réseau neuronal profond dans une infrastructure auto-jouante pour apprendre à réagir à la nouvelle information entrante en adaptant simultanément la séquence d'extraction, la destination des matériaux et l'utilisation des réseaux de traitement à court terme. L'algorithme proposé emploie une recherche arborescente Monte Carlo pour jouer seul au jeu de la planification de production à court terme, afin d'entrainer un agent de réseau neurones profond. Cette thèse propose aussi un algorithme de simulation Monte Carlo pour la mise à jour de l'incertitude liée à la performance de l'équipement. Une application dans un complexe minier de cuivre démontre

une capacité d'adapter la planification à court terme presque en temps réel afin de mieux atteindre les objectifs de production tout en améliorant la production de cuivre et la liquidité.

Les méthodes existantes pour mettre à jour l'incertitude d'approvisionnement des matériaux ne tirent pas profit de la nouvelle information entrante et ne tiennent pas compte des statistiques spatiales d'ordre supérieur. Un nouvel algorithme d'apprentissage par renforcement sans modèle avec agents acteur-critique est donc proposé. Premièrement, l'agent acteur prédit, puis l'agent critique évalue, les attributs mis à jour d'un bloc de minerai, et ces attributs sont ensuite évalués selon une fonction de distribution de probabilité conditionnelle, générée en utilisant des statistiques spatiales d'ordre supérieur puis, finalement, les évaluations sont utilisées pour entrainer les agents. Une application dans une opération minière de cuivre démontre sa faisabilité.

Le cadre d'intelligence artificielle d'auto-apprentissage proposé dans cette thèse peut permettre à un complexe minier d'apprendre et de s'adapter avec la nouvelle information entrante pendant ses opérations, afin de prendre des décisions de planification à court terme plus éclairées. Des recherches futures pourraient améliorer la performance des algorithmes et développer de nouvelles méthodes pour adapter simultanément la planification à long et à court terme.

TABLE OF CONTENTS

xiii

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction

1.1   Overview

A mining complex is an integrated mineral value chain that consists of multiple interrelated activities, starting from mineral deposits and material extraction to a set of sellable products delivered to various customers and/or the spot market. Figure 1-1 presents an example of a mining complex with different components including, but not limited to, suppliers of raw materials (mines/mineral deposits and external sources such as existing stockpiles), earth moving equipment (shovels and trucks), handling facilities (crushers, conveyor belts, re-handling stockpiles, and waste dumps), processing facilities (mineral processing mills and leach pads), tailing facilities, transportation facilities (pipelines, trucks, and, ports), and customers/commodity markets. The primary activities in a mining complex are: (i) -scheduling the extraction of material from available mines, (ii) selecting and allocating equipment selection and allocation to excavate the scheduled material, (iii) determining the destination of extracted material, (iv) blending of materials to supply multiple processing streams (processing and handling facilities), (v) transformation transforming of materials through different processing paths and alternatives (smelter, electrowinning, and refinery), and (vi) transporting of the final sellable products to markets using multiple transportation schemes (pipelines, trucks, and , ports). Supply uncertainty and market uncertainty are major sources of technical risk in a mining complex. They stem from

the uncertainty in the supply of materials from the mineral deposits and existing stockpiles associated with the mining complex, and uncertainty about the price of the different commodities produced, respectively. In addition, equipment performance uncertainty related to the production capabilities of the different equipment is another source of technical risk in an operating mining complex.

Figure 1-1 Example of a mining complex (Goodfellow, 2014)

Supply uncertainty is the uncertainty associated with the geological properties of materials available in the mineral deposits associated with a mining complex, and can be grouped into three categories: continuous properties, categorical properties, and non-additive properties (which can be either continuous or categorical). Continuous properties include grades and deleterious rock properties and rock mass. Categorical properties are material types, lithologies, minezones, mineralization, alterations etc. Continuous and categorical non-additive properties include geometallurgical properties such as hardness, recovery, ore texture, grindability, soft/hard domains, etc.

2

Stochastic simulation methods quantify supply uncertainty by generating a set of equally probable scenarios/realizations that capture the uncertainty and variability of materials. Equipment performance uncertainty is the uncertainty associated with equipment performance properties such as availability, utilization, productivity, breakdowns, cycle time, repair time, and so on. A set of equally probable Monte Carlo stochastic simulations quantify the equipment performance uncertainty.

The strategic or long-term production plan of a mining complex provides annual strategic decisions that maximize the cumulative discounted cash flows (net present value) and meets the different production targets and customer demands. The annual strategic decisions in a mining complex include the determination of the sequence of material extraction, destination policies, processing stream utilization, capital expenditure, and the transportation of products to customers/market decisions. A short-term plan is then generated within the predefined long-term plan to provide daily/weekly/monthly production decisions. The short-term production decisions in a mining complex include defining the daily/weekly/monthly sequence of material extraction, the destination of materials, utilization of processing streams, fleet assignment and allocation, the operating modes of processing streams, transportation schemes, and waste/tailings management decisions.

Simultaneous stochastic optimization of a mining complex aims to simultaneously optimize its several interrelated activities while accounting for related uncertainties, to minimize the deviation from production targets and maximize the value of the integrated mineral value chain. For example, short-term simultaneous stochastic optimization of a mining complex simultaneously optimizes the different short-term decisions, while simultaneously managing supply and equipment

performance uncertainties, to maximize metal production and compliance with the long-term plan and production targets.

A mining complex collects new information during its day-to-day operations according to the short-term plan. This new information can be categorized into two groups, the first being information collected with conventional methods, and the second being information collected using new technologies and advanced sensors. Conventionally collected new information includes geochemical laboratory analysis of blasthole drilling samples to characterize geological properties of blasthole materials, global positioning systems (GPS) installed on the equipment that tracks their movement and location, radio frequency ID (RFID) tags used in blastholes to track the flow of materials in a mining complex, and built-in control units in trucks, shovels, and drilling machines that characterize their fuel consumption, power consumption, brake efficiency, engine torque, temperature, and speed. Control units in crushers, processing mills, and conveyor belts measure their throughput, power consumption, utilization, and so on. New information collected with advanced sensors and new technologies comes from three different types of sensors that differ in terms of what geological properties of materials they measure, and whether such measurements are surface or volumetric. Near-infrared (NIR), short-wavelength infrared (SWIR), mid-wavelength infrared (MWIR), long-wavelength infrared (LWIR), and X-ray diffraction (XRD) sensors generate surface measurements about the categorical geological properties (mineralization) of materials. X-ray fluorescence (XRF) and laser-induced breakdown spectroscopy (LIBS) sensors generate surface measurements of continuous geological properties (grades and deleterious elements) of materials. Dual-energy X-ray transmission (DE-XRT) sensors generate volumetric

measurements of continuous geological properties (grades, deleterious elements, and density) of materials.

The new information generated from conventional technologies and advanced sensors can be used to update the supply uncertainty of materials in the mineral deposits as well as the equipment performance uncertainty. However, the new information, collected spatially with drilling machines at the mineral deposits (spatial data) or collected in time at conveyor belts and processing mills (temporal data), is noisy, therefore uncertain, and referred to as "soft" data. The uncertainty and noise in the incoming new information are attributed to the characteristics of the related sensors that generate indirect measurements (for example, the wavelength of light refracted, transmitted, and absorbed) about the geological properties of materials. Direct measurements, on the other hand, such as those derived from the analysis of drillhole samples in geochemical laboratories, are substantially more precise and known as "hard" data.

The incoming new information, along with the updated supply and equipment performance uncertainties, provides an opportunity for mining complexes to learn to adapt, to make more informed short-term production decisions to better respect their long-term production plan, and better meet their production targets. However, the current combinatorial optimization methods cannot respond to the noisy incoming new information to make real-time decisions. Moreover, the combinatorial optimization methods cannot learn and integrate the noisy incoming new information in the optimization models. Artificial intelligence (AI) agents have shown advanced computational and learning capabilities in recent years to respond quickly to incoming new information, for very complex and challenging planning environments such as

Atari games, traffic control, games of Go and chess, equipment maintenance scheduling, multi-site production, material destinations for mining complexes, and many more. However, unlike the other planning environments mentioned above, a mining complex is substantially more intricate. It has multiple operational and environmental constraints with multiple interlinked components, starting from mineral deposits with spatial characteristics, to the handling of material with trucks, shovels, crushers, conveyor belts, processing mills, tailing facilities, and finally customers/markets that receive the final products. In addition, the description of the different components of a mining complex quantified by their uncertainty models is continuously changing as new information is collected. For example, supply uncertainty of materials in mineral deposits changes as new information regarding quality and quantity of materials is collected from different components of a mining complex with advanced and conventional sensors, and equipment production capabilities change as new information is collected about their performance. A new continuously learning and updating\adapting framework inspired by existing AI algorithms must be developed that can handle the intricacies of a mining complex and can learn, integrate, and respond simultaneously to the incoming new information to adapt the short-term production planning decisions.

1.2   Literature Review

This section presents a review of the literature pertinent to the topics discussed in this thesis to highlight the existing state-of-the-art frameworks and outline their contributions and limitations. Section 1.2.1 discusses the optimization of strategic decisions, or long-term production planning, in a mining complex. Section 1.2.2

describes the quantification of supply uncertainty. Section 1.2.3 details the optimization of the short-term production plan of a mining complex. Section 1.2.4 discusses the quantification of equipment performance uncertainty. Section 1.2.5 provides details about the data collected with sensors in real-time during the day-to-day operations in a mining complex. Section 1.2.6 describes the available frameworks and their limitations for updating the supply and equipment performance uncertainties with the incoming new sensor production data. Section 1.2.7 discusses the state-of-the-art artificial intelligence algorithms and their applications in mining and other planning domains.

1.2.1 Strategic or long-term production planning

Long-term, or strategic planning focuses on optimizing the annual strategic decisions of a mining complex that maximize the life of the asset (net present value) and meet the different production targets. The strategic decisions in a mining complex include defining the yearly sequence of extraction of materials from mines, the destination of extracted materials, the utilization of processing and other facilities, the selection of processing alternatives, the selection of transportation alternatives, and capital investments. Traditionally, these different strategic decisions are optimized in isolation and, more importantly, over estimated orebody models. For instance, the destination of materials is optimized using cut-off grade optimization (Lane, 1984, 1988; Asad and Dimitrakopoulos, 2013; Asad et al., 2016; Del Castillo and Dimitrakopoulos, 2016; Li et al., 2020), the extraction sequences of the materials from the mines are optimized individually using integer/mixed-integer/linear programming models (IP/MILP/LP) (Johnson, 1969; Gershon, 1983; Barbaro and Ramani, 1986;

7

Underwood and Tolwinski, 1998; Hustrulid et al., 2013), and processing stream utilization decisions are optimized separately with another optimization model, e.g. Farzanegan and Vahidipour (2009). This separate stepwise optimization of multiple interrelated components of a mining complex generates sub-optimal long-term production plans.

Earlier attempts at long-term simultaneous deterministic optimization of multiple components of mining complexes include Hoerger et al. (1999), Stone et al. (2007), Whittle (2007), Whittle and Whittle (2007), and Whittle (2010). Hoerger et al. (1999) presented a model that optimizes mining decisions (timing of open-pit layback and underground stope development, capital expenditures and mining rates), destination decisions (mine to plant, mine to stockpile, and stockpile to plant material flow), and processing decisions (timing of plant startup and shutdown, capital expenditures, ore processing rates) simultaneously. The model has material flow decisions as linear variables, and plant startup and shutdown, extraction sequencing, and capital expenditure decisions as integer variables. Stone et al. (2007) presented a mine planning optimization tool, "Blasor," that optimizes long-term mining, destination and blending decisions to maximize the discounted cash flows (NPV) while meeting production targets. Blasor is limited in the sense that: (i) it only optimizes the proportion of material extracted annually from each of the multiple pits (mining decision), rather than generating an extraction sequence for multiple pits, (ii) it aggregates mining blocks into aggregation units/panels to reduce the computational requirement to solve the model, and (iii) it ignores the supply uncertainty of mineral deposits. The stochastic version of "Blasor" that considers geological supply uncertainty is shown in Menabde et al. (2007). The stochastic version of "Blasor"

8

optimizes the long-term mining, destination, and blending decisions to maximize the expected discounted cash flow under supply uncertainty. In addition, the maximum mining and processing rate constraints need to be respected over all the scenarios of supply uncertainty. However, with this approach, if any simulation violates these constraints, then the solution for the optimization model becomes unfeasible. Whittle (2007) presented a model for optimizing several components of mining complexes simultaneously and termed it "global asset optimization." However, the model uses a sequential optimization approach, which first defines the production schedules and then locally optimizes blending and processing stream decisions. Therefore, although the model considers multiple components of mining complexes, it does not optimize them all simultaneously. In addition, it aggregates mining blocks into panels and parcels to reduce computations requirements, and disregards supply uncertainty. Whittle and Whittle (2007) used the global asset optimization model for two case studies: (i) five pits with a total of 37 material types, five concentrators and multiple stockpiles that produce 14 products and (ii) 130 pits with five material types, concentrators, and stockpiles that produce two products. Whittle (2010, 2014) presented the ProberB and ProberC algorithms used in global asset optimization to solve mining complexes. The algorithm repeatedly creates a random feasible solution (a solution that satisfies mining constraints) and then finds the nearest local maximum solution using a linear programming model. An integrated approach was presented by Pimentel et al. (2010) to address the simultaneous optimization of mining complexes and possible solution strategies. Place et al. (2018) presented the advanced deterministic simultaneous optimization model (SIMO) that uses the ProberB algorithm to optimize the extraction sequences, stockpiling strategies, cut-off grade

9

policies, and blending strategies of multiple mines. Similar to Whittle's global asset optimization model (Whittle, 2007), the algorithm aggregates the blocks into pushbacks, panels, and blend bins, converts the non-linear simultaneous optimization model into a linear model using approximation techniques, generates a random feasible schedule, finds the nearest local maximum by solving the linear model and, finally, uses a method analogous to the Monte-Carlo method to start from the different initial feasible solutions to find a global maximum. The deterministic simultaneous optimization models discussed above have several limitations. For example, they all (i) consider block economic value (defined using a pre-optimized cut-off grade policy) in the optimization model, (ii) combine mining blocks into aggregate units such as parcels and panels to overcome the computational requirements of solving such large optimization models of mining complexes, (iii) do not jointly optimize all the relevant long-term production planning decisions in a mining complex, and (iv) do not account for uncertainty related to the supply of materials and to commodity prices in the optimization models.

Recent advances in the long-term simultaneous stochastic optimization of multiple components in a mining complex are based on two-stage stochastic optimization techniques developed in the last decade that account for supply uncertainty, consider a single mining operation, and focus on the economic value of blocks (Dimitrakopoulos, 2011). The stochastic approaches developed in the past can be grouped into three different categories. The first approach involves models that minimize deviations from ore and waste production targets and are based on a simulated annealing algorithm (Godoy, 2002; Godoy and Dimitrakopoulos, 2004; Leite and Dimitrakopoulos, 2007; Albor Consuegra and Dimitrakopoulos, 2009). The

second approach involves models that maximize the expected net present value (NPV) while minimizing deviations from production targets, and are based on stochastic integer programming (Menabde et al., 2007; Ramazan and Dimitrakopoulos, 2007, 2013). The third approach illustrated in Boland et al. (2008) uses a multistage programming approach that considers scenario-dependent mining and processing decisions to provide a set of policies (mining and processing decisions) to follow according to the scenario observed during the progress of the extraction. However, this approach is not operationally realistic because it does not provide a single extraction sequence to follow and, moreover, it is not possible to differentiate and identify which scenario is observed during the advancement of the mining operation.

The recent models of long-term simultaneous stochastic optimization of multiple -product mining complexes optimize the multiple mines, destinations, and processing streams in a single optimization model which focuses on the value of the products sold, considers supply and market uncertainty, and accounts for the non-linear interactions that happen in a mineral value chain. Goodfellow and Dimitrakopoulos (2016, 2017) proposed a model (first and the second part of the objective function represented in Eq. 1.1) for the long-term simultaneous stochastic optimization of mining complexes that integrates the extraction sequence, cluster-based destination policies, and processing streams utilization decisions. The first part of the objective function maximizes the expected value of products, minus any costs incurred to generate products over the planning horizon and under supply uncertainty.

$$\max \frac{1}{|\mathbb{S}|} \underbrace{\sum_{s\in\mathbb{S}} \sum_{t\in\mathbb{T}} \sum_{i\in\mathbb{L}} \sum_{a\in\mathbb{H}} p_{a,i,t} \cdot v_{a,i,t,s}}_{\text{Part I}}$$

$$-\frac{1}{|\mathbb{S}|} \underbrace{\sum_{s\in\mathbb{S}} \sum_{t\in\mathbb{T}} \sum_{i\in\mathbb{L}} \sum_{a\in\mathbb{H}} c^+_{a,i,t} \cdot d^+_{a,i,t,s} + c^-_{a,i,t} \cdot d^-_{a,i,t,s}}_{\text{Part II}} - \underbrace{\sum_{t\in\mathbb{T}} \sum_{k\in\mathbb{K}} p_{k,t} \cdot w_{k,t}}_{\text{Part III}} \quad (1.1)$$

The second part of the objective function minimizes the deviations from the upper and lower production capacities. The third part of the objective function accounts for the expenses that are incurred by exercising a capital investment decision $k \in \mathbb{K}$ at time $t \in \mathbb{T}$. Here, $\mathbb{K}, \mathbb{T}, \mathbb{L},$ and $\mathbb{S}$ define the set of capital investments, time periods, locations, and supply uncertainty scenarios, respectively. The quantity of attribute $a \in \mathbb{H}$ at location $i \in \mathbb{L}$ at time $t \in \mathbb{T}$ under supply uncertainty scenario $s \in \mathbb{S}$ is defined by $v_{a,i,t,s}$. The set, $\mathbb{H}$, represents the set of rock properties such as the quantity of rock mass, copper mass, gold mass, etc., which are additive. The discounted price and costs per unit of attribute h at location i at period t are denoted by $p_{a,i,t}$. The variables $d^+_{a,i,t,s}$ and $d^-_{a,i,t,s}$ calculate the deviation from upper and lower production capacities, respectively, for the quantity of attribute h at location i at time t in scenario s. The constants $c^+_{a,i,t}$ and $c^-_{a,i,t}$ define the discounted penalty cost associated with deviating from the upper and lower production capacities, respectively, for the quantity of attribute h at location i at time t in scenario s. The decision variable $w_{k,t}$ represents the quantity of investment k exercised with a unit discounted cost of $p_{k,t}$ at time t. The model maximizes the value of products, $p_{a,i,t} \cdot v_{a,i,t,s}$, generated in a mining complex minus any cost incurred to generate the products, instead of block economic values. Accounting for the value of products allows this one single optimization

12

model to optimize all the relevant long-term production planning decisions simultaneously, and to account for any non-linear relationships that exist within a mining complex. The extraction sequence decisions define the period in which blocks are extracted from multiple mines. Cluster-based destination policies first generate clusters of materials based on the multivariate properties of materials in the mines (such as copper, gold, silver, arsenic etc.), and then decide the destination of the clusters in the optimization model. Therefore, the cluster definition is scenario-independent, but block destination is scenario-dependent because of the block membership the changes to different clusters in different supply uncertainty scenarios. Processing stream utilization decisions define the proportion of materials sent from one location to another in a mining complex while preserving mass balancing of materials. Results from a copper-gold operation indicate large deviations from capacity targets of sulphide leach pad and sulphide mill (40% and 31%) when a conventional mine plan is tested under supply uncertainty of mineral deposits. In the case study, simultaneous stochastic optimization model manages the technical risk associated with the geological supply uncertainty better and reduces the average deviation in capacity target for sulphide leach pad to 10% and sulphide mill to less than 1% in the initial 10 years to 12% near the end of life-of-mine, while maximizing the cumulative discounted value of the business to 22.6 % compared to conventional plan. Goodfellow and Dimitrakopoulos (2017) used the proposed model at a Nickel-laterite mining complex to simultaneously optimizes the cluster-based destination policies and processing stream utilization and achieves a 3% higher value with an average deviation of less than 1% for capacity and blending targets compared to results from deterministic optimization on estimated mineral deposits using the

13

proposed method. Saliba and Dimitrakopoulos (2019a) used this model at a gold mining complex with supply and market uncertainty. In the case study, the model can simultaneously optimize all the components to reduce operating complexity by utilizing only 12 stockpiles compared to 38 in the current mine plan. Additionally, the model can account for market uncertainty by adapting the schedule to mine and process more materials during periods of the elevated price while being conservative during downside exposure of prices. Levinson and Dimitrakopoulos (2019) used this model at a gold mining complex under supply uncertainty to manage and reduce waste production. In the case study, the model can balance the requirement of the processing facility and potentially acid-generating waste to satisfy the environment, permitting, and processing limits while increasing the NPV by 6%.

Goodfellow and Dimitrakopoulos (2015) and Farmer (2016) proposed a long-term simultaneous stochastic optimization model of mining complexes that integrates extraction sequence, destination policies, processing stream utilization and capital expenditure decisions (complete objective function represented in 1.1). The model includes constraints that account for changes in the mining and processing rates with capital investment decisions. A case study of the proposed model at a copper mining complex shows 5.4 % higher NPV, higher utilization of processing mill capacities, and better management of mining investment decisions (truck and shovel), compared to the deterministic design. Farmer (2016) accounts for both processing (mill) and mining (truck and shovel) capital investment decisions in the proposed model. A case study at a copper-gold mining complex shows an increase of 12% in NPV compared to the deterministic design. Farmer (2016) also proposed an approach to account for market uncertainty. First, the extraction sequence, capital expenditure, and cluster-

based destination policies are optimized accounting for supply uncertainty, such decisions are then fixed, and finally, the processing stream utilization decisions are optimized, accounting for market uncertainty. Saliba and Dimitrakopoulos (2019b) used this model for the simultaneous optimization of a gold mining complex with tailings management under supply uncertainty. In the case study, the model can find the optimal quantity and timing of tailing facility expansion to enlarge mine footprint, generating 14% and 4% of higher gold ounces and NPV respectively compared to a model without tailing management.

Montiel and Dimitrakopoulos (2015, 2017, 2018) proposed a model for long-term simultaneous stochastic optimization of mining complexes that incorporates extraction sequence, block destination policies, processing mill operating modes, transportation alternatives, and processing stream utilization decisions. Block destination policies define the destination of blocks as scenario-independent decisions over the supply uncertainty scenarios. Processing mill operating modes decisions defines how to process materials at the processing mill (fine or coarse grinding). Transportation alternative decisions define what quantity of processed materials to transport to customers with trucks and pipelines. Results from a multi-pit copper operation with stringent blending requirements indicated that conventional mine plans generated using commercial mine planning software perform adversely when tested under supply uncertainty. Significant and impractical deviation in terms of capacity (40% and 30%) and blending targets (11% and 22%) were observed. In the given case study, the simultaneous stochastic optimization model minimizes such large deviations from capacity targets to 1% and 3% and blending targets to 0.7% and 1.2% for two mills, while improving the value of the business to 5% compared to the

conventional mine plan. Montiel and Dimitrakopoulos (2017) used this model at a copper mining operation indicated a significant and impractical deviation in the capacity target (45%) when the conventional mine plan is tested under supply uncertainty. The proposed method minimized the deviation to less than 1% while increasing the value of the business to 30%, as compared to the conventional mine plan. Montiel et al. (2016) used this model at a mining complex comprised of multiple open-pits, underground operations, and destinations under geological supply uncertainty. Results from the copper-gold operation with stringent blending targets indicated large and impractical deviations in the capacity target for open-pit (60%) and underground mine (40%) and blending targets for the mill (SS/CO3 – 50%, SS – 45%, CO3- 25%) when tested under supply uncertainty. The proposed method minimizes the deviation from capacity targets to 1%, and 15% and blending targets to 20%, 15%, and 10%, along with a 14% increase in the value of business compared to the conventional mine plan. Montiel and Dimitrakopoulos (2018) used this model to optimize long-term production plan of Newmont's Nevada mining complex that consisted of two open pit mines, three external sources of concentrate materials, an oxide mill, an autoclave, a heap leach, a waste dump, multiple re-handling sulfide and oxide stockpiles. An adaption technique is then applied to adapt the simultaneous stochastic long-term production plan to be operational given the existing infrastructure, mining constraints, access and mining equipment available at the mining complex. The mine's plan is tested under the supply uncertainty of materials and showed a 4% decrease in forecasted recoverable gold, 6% decrease in NPV, and large and impractical deviations from blending and capacity requirements. The adapted simultaneous stochastic long-term production plan increases the forecasted

recovered gold by 6%, NPV by 7%, and satisfy the blending and capacity requirements. In addition, the adapted simultaneous stochastic long-term is mineable given the constraint of existing shapes from previously mined-out areas and existing infrastructure.

Simultaneous stochastic optimization models of mining complexes are large combinatorial optimization models with 10^{ths} of millions of binary decision variables, which requires substantial computational time to solve using general-purpose commercial solvers such as CPLEX. Metaheuristic algorithms based on methods such as multiple neighbourhood simulated annealing (Goodfellow and Dimitrakopoulos, 2016; Montiel and Dimitrakopoulos, 2017), Tabu search (Lamghari and Dimitrakopoulos, 2012), variable neighbourhood descent (Lamghari et al., 2014), hybridization of linear programming and variable neighbourhood descent (Lamghari et al., 2015), progressive hedging (Lamghari and Dimitrakopoulos, 2016a), network flow (Lamghari and Dimitrakopoulos, 2016b), ant colony optimization (Gilani and Sattarvand, 2016), and hyper-heuristic (Lamghari and Dimitrakopoulos, 2018), were proven efficient for solving such large optimization model of mining complexes. A review of the metaheuristics algorithms for optimizing long-term plan of mining complexes is discussed in Lamghari and Dimitrakopoulos (2017) and Franco-Sepúlveda et al. (2019).

Extensions of simultaneous stochastic optimization models are presented in Del Castillo and Dimitrakopoulos (2019), Del Castillo (2018) and Zhang and Dimitrakopoulos (2017). Del Castillo and Dimitrakopoulos (2019) and Del Castillo (2018) proposed a dynamic simultaneous stochastic optimization model that integrates dynamic mining and processing capacities (capital investment), extraction

sequence, processing stream utilization and destination policy decisions in the long-term simultaneous stochastic optimization of mining complexes under geological supply uncertainty. The method first fixes extraction and capital investment decisions as first-stage for period t and second-stage for others and optimizes the model to generate scenario-dependent mine plans for periods other than t. If the number of scenario-dependent mine plans is higher than a specified threshold value, the plan branches for that decision and the whole model is reoptimized fixing period t and $t +$ 1 as the first-stage and others as the second-stage. The model, therefore, compromises between risk management and flexibility, providing a fixed plan for initial years, but at the same time allowing dynamic long-term decision change. A case study at a copper-gold mining complex indicated that the simultaneous stochastic optimization model with dynamic capital investment decisions generates a 3 % higher value compared to a model where capital investment decisions are set as first-stage.

Zhang and Dimitrakopoulos (2017) proposed a long-term dynamic material-value-based Benders decomposition method to decompose the simultaneous stochastic optimization model. The method decomposes the model into a master model with mine production scheduling optimization under supply uncertainty and a sub-model with material flow optimization under market uncertainty. The solution of the two models is then iteratively synchronized to determine the optimal mine production schedule accounting for the downstream congestions and market fluctuations. Results from a copper mining complex indicated that (i) the life of the two mines increased from 8 to 12 years and from 14 to 20 years, (ii) the stockpile quantity is reduced by 33 % for both of the mines, (iii) investment in the processing

18

capacity is reduced by 13 % at the beginning of planning horizon, (iv) utilization of processing capacity is increased from 74 % to 88 %, (v) increase in the NPV of the mineral value chain by 6 %, as compared to the local optimization approach where the mine production schedule is optimized for an individual mine without considering the downstream material flow.

Supply uncertainty related to non-additive geometallurgical properties such as texture, granularity, hardness, recovery, power consumption, etc. influences the performance of the processing mills in a mining complex. Geometallurgical properties of materials affect the mine planning decisions (Coward et al., 2009; Dunham and Coward, 2011; Macfarlane and Williams, 2014). Macfarlane and Williams (2014) suggested that using the geometallurgical properties in the long-term mine planning optimization will help to maintain a consistent geometallurgical and grade feed of materials, which can then generate better NPV. Williams and Richardson (2004) proposed a geometallurgical mapping approach that uses ore characterization procedure to map geometallurgical response of mining blocks into the orebody models for forecasting mining block recoveries. With this method, the future cash flows of the mining operation can be generated while incorporating the metallurgical response of the individual mining blocks in the mine planning models.

Kumral (2011) proposed an approach that first associates mining blocks with their recoveries, throughputs and processing costs using a regression model of historical processing mill data. It then uses two models to solve the mine production scheduling optimization, first with a two-stage stochastic model and second with maximin model. Navarra et al. (2018) proposed a method that combines mineralogy, liberation, texture, and mineral chemistry data of each mining block to compute a

geometallurgical factor that then defines the recovery, processing cost, and processing rate of mining blocks in the supply uncertainty scenarios for different processing modes. It then used the supply uncertainty scenarios within a two-stage stochastic optimization to decides the extraction sequence as a first-stage decision and then decides where to process the materials as a second-stage scenario dependent decision which looks at the recovery, processing cost and processing rate of the blocks in the given scenario to decides the best processing operating mode. Morales et al. (2019) proposed a two-stage stochastic optimization model that integrates recoveries and throughput (time required to process materials) of mining blocks in mine production scheduling optimization. The model accounts for uncertainty in individual block processing time and recovery and then maximizes the NPV to meet the constraints with the maximum total available annual processing time. However, all the methods assume that each mining block is processed independently, and there is no blending of materials. However, in a mining complex, the materials are blended and processed together, which affects their response in the processing mills.

The simultaneous stochastic optimization models mentioned above optimizes the relevant long-term production planning decisions in mining complex in one single optimization model, but do not account for supply uncertainty related to geometallurgical properties of materials and do not provide links to the short-term production plan.

1.2.2 Quantification of supply uncertainty

Supply uncertainty is a major source of technical risk in long-term planning. Studies have shown that long-term plan based on estimated orebody model of mineral

deposits is not realized when tested under supply uncertainty of materials (Ravenscroft, 1992; Dowd, 1994, 1997; Dimitrakopoulos et al., 2002; Godoy, 2002; Ramazan and Dimitrakopoulos, 2007, 2013; Goodfellow and Dimitrakopoulos, 2017; Montiel and Dimitrakopoulos, 2018; Mai et al., 2019). Stochastic simulation methods are efficient at quantifying the supply uncertainty by generating a set of equally probable representation of mineral deposits, referred to as stochastic orebody models/simulations. Stochastic orebody simulations reproduce the statistics of available drill hole information and the local variability of materials, compared to estimated orebody models that generate a smooth representation of mineral deposits by smoothing the distribution of extreme values and local variability of materials.

1.2.2.1 *Continuous and categorical geological properties*

In the 1970s, turning-bands method (Journel, 1974; Journel and Huijbregts, 1978) was proposed to generate stochastic simulations of mineral deposits; however, the method had a loss of accuracy due to approximation and difficulties in reflecting anisotropic covariances. Later, a sequential simulation framework (Deutsch and Journel, 1992; Journel, 1994) was proposed that overcomes such limitation and is a well-known method up to now for quantifying geological supply uncertainty by generating multiple stochastic orebody simulations. It is based on the decomposition of the multivariate probability density function (PDF) of a stationary and ergodic random process $Z(u)$, $u \in R^n$ to the product of univariate posterior distribution functions and then sequentially sampling the posterior probability density functions to generate simulated value for a point. For instance, let $f(z_1, z_2)$ be the probability distribution function with a bivariate process $Z = \{Z_1, Z_2\}$ at $u = \{u_1, u_2\}$, generating

21

realization begins with product decomposition of univariate posterior PDF function, $f(z_1, z_2) = f(z_1) \cdot f(z_2|z_1)$, and then sampling a value z_1 based on $f(z_1)$; z_2 is drawn based on $f(z_1, z_2)$. For a N-variate (simulation grid size) process considering stationary Gaussian random field $Z(u_i), u_i \in R^n, i \in \{1, N\}$, and a set of available conditioning data $d_n = \{z(u_\alpha), \alpha \in \{1, n\}\}$, then the N-variate probability distribution posterior to the data set d_n such that $\beta_1 = \{d_n, Z_1\}, \ldots, \beta_N = \{d_n, Z_1 \ldots Z_N\}$ is given by Equation 1.2 with a conditional cumulative distribution function (CCDF) equal to the product of N single-variate posterior density functions (CPDF) given by Equation 1.3.

$$F(u_1, \ldots, u_N; z_1, \ldots, z_N | d_n) = P(Z(u_1) \leq z_1, \ldots, Z(u_N) \leq z_N | d_n) \qquad (1.2)$$

$$f(u_1 \ldots u_N; z_1 \ldots z_N | d_n) = \prod_{i=1}^{N} f(u_i; z_i | \beta_{i-1}) \qquad (1.3)$$

When $Z(u)$ is Gaussian, the method is termed as sequential Gaussian simulation (SGS) (Deutsch and Journel, 1992; Journel, 1994; Goovaerts, 1997; Remy, 2005; Remy et al., 2009; Rossi and Deutsch, 2013). Dimitrakopoulos and Luo (2004) introduced group sequential Gaussian simulation (GSGS) to improve the computational efficiency of SGS. In GSGS, the simulation grid, N, is divided into a set of groups of size v, and the nodes within the group share the same neighbourhood with a maximum number of conditional data, v_{max}. A random path is then selected to visit each group, the nodes inside this group is simulated using lower-upper decomposition (LU) (Davis, 1987) and added as the conditioning data, the process continues until all groups are visited. This reduces the computational cost of GSGS because the nodes inside the groups are simulated simultaneously compared to

visiting each node to solve distinct, simple kriging equations in SGS. A critical problem of the SGS and GSGS methods is that the simulation is performed at a point-support scale, as seen with Eq. 1.3, which needs to be re-blocked into the block scale. The block support corresponds to selective mining units that are used for planning purposes. Re-blocking of point-support is often generated by approximating the block value through a linear average of contained point values. Such linear averaging introduces problems with connectivity of high-grade values, which are essential for planning purposes and require substantial time and memory to store the point-support simulated values. Godoy (2002) introduced the method direct block simulation (DBSIM) that simulates directly at block support scale. The method uses the block size as selective mining unit dimension and then subdivides the block into internal nodes. A random path is selected to visit each block, and the internal nodes are simulated simultaneously using LU decomposition, the internal simulated nodes are averaged and then discarded to generate the simulated block value which is used as new conditioning data until all the blocks are visited. The method accounts for point-point, block-point, and block-block covariances during the simulation of internal nodes.

Mineral deposits frequently consist of multiple spatially correlated elements/properties of interest. Stochastic orebody simulations should preserve such spatial correlation of multiple elements in the mineral deposits. In the past, principal component analysis (PCA) has been used to de-correlate the collocated correlated variables (Davis and Greenes, 1983; David, 1988); however, the de-correlation with PCA only guaranteed decorrelation of the zero-lag distance covariance matrices ignoring the spatial correlation of the variables at non-zero lag distances. Switzer and

Green (1984) proposed an alternative method, called minimum/maximum autocorrelation factor (MAF), that uses two PCA decomposition, to transform a set of spatial collocated correlated elements into independent factors at all the lag distance, provided there are up to two structures in the linear model of coregionalization. MAF performs two sequential eigenvalue decomposition of the covariance matrix to decorrelate the variables into independent MAF factors. For this case, cross-variogram matrix, $\Gamma(h)$, at a lag distance of h is given by Eq. 1.4.

$$\Gamma(h) = B_1\gamma_1(h) + (B - B_1)\gamma_2(h) \tag{1.4}$$

$$Y(u) = \Lambda^{1/2}Q^TZ(u) = AZ(u) \tag{1.5}$$

$$F(u) = Q_1^TY(u) = Q_1^T\Lambda^{1/2}\, Q^TZ(u) \tag{1.6}$$

B_1 is the cross-covariance matrix for the first structure, B is the variance-covariance matrix, Λ is a diagonal matrix in decreasing order of eigenvalues of B, and $\gamma_1(h)$ and $\gamma_2(h)$ are the unit variograms for elements 1 and 2, respectively. The first spectral decomposition gives $B = Q\Lambda Q^T$ where Q is the orthogonal matrix of eigenvectors, then PCA factors with respect to B is given by Eq. 1.5. The second spectral decomposition gives $AB_1A^T = Q_1\Lambda_1Q_1^T$, where Q_1 is an orthogonal matrix of the eigenvectors. Then the MAF factors are expressed using Eq. 1.6. These MAF factors can then be simulated independently and back-transformed into original data space to maintain the spatial correlation in the orebody simulations. Desbarats and Dimitrakopoulos (2000) applied the MAF method for generating stochastic simulations of regionalized pore size distribution and observed the excellent performance of the method for retaining the spatial correlation of correlated variables in the simulations. Bandarian et al. (2008) proposed a direct minimum/maximum

24

autocorrelation factor that transforms the multiple correlated elements into MAF factors without normal score transformation and further presented the efficiency of the method for a multivariate soil dataset. Boucher and Dimitrakopoulos (2009) proposed direct block simulation with minimum/maximum autocorrelation factor (DBMAFSIM) method for generating stochastic simulations, that combines the DBSIM and MAF technique to efficiently simulate large deposits with multiple correlated elements at block support scale, further utilized to simulate iron ore deposit with multiple correlated elements in Boucher and Dimitrakopoulos (2012). The Gaussian simulation methods described above require the random variable to be in Gaussian space. Therefore, in practice, a normal score (Gaussian transformation) is performed to convert original data into Gaussian space before the simulation method begins. The aforementioned two-point simulation methods rely on variogram models to infer pairwise second-order spatial statistics, which is not sufficient to characterize complex spatial non-linear patterns and curvilinear geological structures that are typically found in a mineral deposit (Journel, 2005, 2007).

Figure 1-2 Limitations of two-point simulation methods for the simulation of complex geological structures (Journel, 2007)

Figure 1-2 represents the limitation of two-point spatial statistics methods where three significantly different geological structures present the same variogram model. The figure shows that three different geological patterns share the same variogram models and are therefore hard to differentiate if only two-point statistics (variograms) inferred from the drillhole samples are considered during the simulation process. In contrast, multiple-point simulations methods that consider multiple points simultaneously and uses multiple-point statistics are efficient at modelling and reproducing the complex and non-linear geological structures and patterns. Multiple-point simulation method visits the simulation node in the simulation grid in a random path, then finds n closest conditioning data points within a neighbourhood. The geometry of the n points around the simulation (central) node is defined by a spatial distance-vector called template. The conditioning values of the n points in the

26

template is called the conditioning data event. It then searches for replicates that match the template in the training image (TI) to infer the multiple-point spatial statistics to generate the probability density function. The probability density function is sampled to generate the simulated value for the central node and then added as new conditioning data in the followed simulations. TI is a three-dimensional representation of geological structures generated using sparse drill hole information, and therefore is subjective to the perception of a geological modeller.

Guardiano and Srivastava (1993) introduced a multiple-point simulation (MPS) pixel-based method called ENESIM further developed in Strebelle (2002) to a computationally manageable multiple-point statistics method called single normal equation simulation (SNESIM). ENSEIM algorithm first transforms the original data into indicator/categorical codes using Eq. 1.7 and then the corresponding indicator random function represented by Eq. 1.8 is used to characterize the random field.

$$i(u; z_k) = \begin{cases} 1 \ if \ z(u) \leq z_k \\ 0, otherwise \end{cases} \ \forall \ k \in 1, \dots, K \tag{1.7}$$

$$I(u; z_k) = \begin{cases} 1 \ if \ Z(u) \leq z_k \\ 0, otherwise \end{cases} \ \forall \ k \in 1, \dots, K \tag{1.8}$$

Here, z_k are the threshold value used to divide the original data into $K + 1$ classes. At each simulation node $Z(u_o)$ visited in a random path, a template ζ defining the geometry is formed by finding the n closest conditioning data points. Let $d_n = \{Z(u_1) = z_1, \dots, Z(u_n) = z_n\}$, where $z_i \in \{1, \dots, K\}$, represent the data event values of the template. The TI is then searched to find replicates of the template ζ. A system of simple kriging equations is solved using single normal equations to derive the conditional probability distribution represented by Eq.1.9.

$$Prob[Z(u_o) = z_k | d_n] = \frac{Prob[Z(u_o) = z_k, d_n]}{Prob(d_n)} \qquad (1.9)$$

Here, $Prob[Z(u_o) = z_k, d_n]$ represent the probability that the central node in the replicates of the template ζ is z_k. $Prob(d_n)$ represent the probability of occurrence of the template, ζ, in the TI. The conditional probability distribution is sampled and added to the conditioning data. The method continues until all the nodes in the simulation grid are visited. The conditional probability distribution function now captures multiple points statistics. However, the algorithm has a high computational cost because of the scanning of TI for replicates every time a simulation node is visited and simulated. SNESIM algorithm overcomes this limitation by storing all the possible template configuration and its statistics by scanning the TI only once as a preprocessing step. The tree is then used to derive the statistics of the template when a node is visited and simulated. This method is further improved to overcome the memory requirement of creating the tree in Straubhaar et al. (2011) to include list-based storage of patterns by direct sampling of the patterns (Mariethoz and Renard, 2010) from training images. The direct sampling method does not create a tree; instead, for every data event, it randomly picks a replicate from the TI and then uses a distance function to define the similarity between replicates and the data event. If the replicate is similar, then the simulation node is assigned as the central node value of the replicate; otherwise, it picks another replicate. Strebelle and Cavelius (2014) improved the SNESIM algorithm and proposed two methods to reduce the data search neighbourhood (i) multiple-grid approach introducing additional intermediary sub-grids and (ii) new search neighbourhood to preferentially include previously simulated nodes. Goodfellow et al. (2012) used the SNESIM method for simulation

of categorical geological property minezones and the DBMAFSIM method for simulation of correlated continuous geological properties nickel, copper, gold, platinum and palladium. The case study showed that the SNESIM method generates simulations/realizations that reproduces the spatial statistics of the TI, since the method relies on a TI to infer spatial statistics, but do not reproduces the spatial statistics of the drillhole data. Other MPS pattern-based methods for simulation of categorical and continuous geological properties are SimPat (Arpat and Caers, 2007), FilterSim (Zhang et al., 2006), WaveSim (Chatterjee et al., 2012, 2016; Chatterjee and Dimitrakopoulos, 2012), and are based on distance-based similarity measurement functions. Avalos and Ortiz (2019) presented a method for simulation of categorical geological properties using a convolutional neural network (CNN). The CNN is trained by extracting the template and its data event values from a TI. The simulation grid is then visited in a random path, the data event formed for each simulation node, the date event with the template is fed to the CNN that outputs the value of the central. Multiple simulations are generated by changing the random path to visit the nodes in the simulation grid.

Pixel and pattern-based MPS methods described earlier only accounts for multiple-point statistics in an ad-hoc way by either only accounting for the occurrence of a data event in the TI or by matching the data event with the replicate from TI through distance-based similarity measurements. These methods do not account for low and high-order spatial statistics and the relations between the different values in the data event. In addition, they rely on TI to infer the statistics and therefore reproduces the statistics of TI. However, a TI is hard to acquire, and moreover, is subjective and can have spatial statistics different from the conditioning data leading

to simulations that conflict with the spatial statistics of the conditioning data. Mustapha and Dimitrakopoulos, (2010a) proposed high-order stochastic simulation (HOSIM) that uses spatial statistics (spatial cumulant and moments) for simulation of complex geological structures and non-linear patterns. Spatial moments for a template ζ with data event d_n is defined using Eq. 1.10

$$Mom_Z^\zeta(w_1, \dots, w_n) = E\left[h_1, \dots, h_n; Z_1^{w_1}, \dots, Z_n^{w_n}\right] \qquad (1.10)$$

Here, $h_i, \forall i \in n$ defines the distance vector that represents the geometry of template ζ and $w_i, \forall i \in n$ defines the order of moments for each node in data event $Z(u_i), \forall i \in n$. The simulation algorithm still follows the sequential simulation framework. Templates are defined as prior to scan the TI and conditioning data for replicates and compute spatial cumulants. The spatial cumulants, along with its templates, are then stored in a tree. The nodes in the simulation grid are visited in a random path to find the conditioning data event. The tree is used to infer the spatial cumulants of the data event and Legendre polynomials are used to approximate the CPDF build with spatial cumulants. A value is sampled from the CPDF and added to the conditioning data. However, defining templates as a priori to avoid computation costs is not efficient because the quality of simulation then becomes dependent on the set of templates used. This method is further used in Mustapha and Dimitrakopoulos (2010b, 2011), Mustapha et al. (2011), and is parallelized in Li et al. (2014). Yao et al. (2018) presented a computational model for high-order simulation that does not require a prior definition of templates and computation of moments or cumulants and reduces the computational requirements of the HOSIM method significantly. The method improves the high-order simulation method by approximating the CPDF using

spatial Legendre polynomials with location-dependent templates. Minniakhmetov et al. (2018) improved and stabilized the performance of the high-order simulation method by using Legendre-like orthogonal splines to approximate the CPDF. The proposed Legendre-like splines functions were able to approximate better the CPDF. de Carvalho et al. (2019) presented an extension of Legendre-like orthogonal spline high-order simulation method for simulations at the block support scale. The method computes the CCDF by using the high-order spatial statistics of point to block relation inferred from a training image. Minniakhmetov and Dimitrakopoulos (2017a) proposed a joint high-order simulation method for simulation of spatially correlated continuous geological properties. The method first orthogonalizes the correlated elements into independent factors using the diagonal domination condition of high-order cumulants; then, the independent factors are simulated independently using high-order univariate simulation method based on high-order spatial cumulants and Legendre polynomials; finally, the simulated factors are back-transformed to generate simulations that respect low-and-high-order spatial statistics. Minniakhmetov and Dimitrakopoulos (2017b) presented a high-order data-driven simulations method that does not require training images for simulation of categorial geological properties in mineral deposits. The proposed method used high-order spatial indicator moments and advanced recursive B-splines to approximate the CPDF. The method further showed that the high-order indicator moments are related to lower order through a boundary condition. de Carvalho and Dimitrakopoulos (2019) showed the effect of using high-order simulations compared to two-point SGS simulations in the mine production scheduling optimization. High-order simulations reproduce the spatial connectivity of high and low grades and result in a more informed long-term mine

31

production plan with 5% more NPV compared to mine production plans based on simulations generated using SGS method.

1.2.2.2 *Non-additive geometallurgical properties*

Simulations of geometallurgical properties of materials are hard to generate and even harder to integrate into the long-term planning optimization models due to their non-additive properties. There are three major categories of approaches available in the literature for modelling the geometallurgical properties of materials, first that models the geometallurgical properties as a function of other rock properties, second that generates domains/clusters of materials with similar geometallurgical properties, and third that models the geometallurgical properties by generating orebody simulations using Gaussian simulation methods.

Alruiz et al. (2009) used a model to simulate the response of the ball mill circuit with different hardness indexes. The simulation results are then used to fit a simple power model that predicts the power consumption of the ball mill circuit. The hardness values of materials can then be related to the ball mill power consumption and throughput. Coward et al. (2009) proposed an approach to divide the geometallurgical properties in the primary and response framework, where primary represents the geometallurgical properties such as density, colour, grain size, alternation, etc. that can be directly measured and response that represents the non-linear geometallurgical properties such as throughput, grindability, recovery, size distribution etc. of the materials. Boisvert et al. (2013) developed a two-stage linear regression model to predict different geometallurgical properties of interest from 204 different input variables and achieved a Pearson correlation coefficient between 0.65

32

and 0.9. However, the methodology was based on different grouping variables based on consultation with a geologist to simplify the regression model input and overcome the problem of sparse sampling and missing information. Sepúlveda et al. (2017) proposed a projection pursuit method to model geometallurgical properties. Projection pursuit is a statistical modelling technique in which data from several variables (grade, tonnage, which are additive properties) are projected onto a direction that optimizes the fit of the regression models for the prediction of geometallurgical properties. Lishchuk et al. (2019) studied different machine learning algorithms, namely, elastic net, support vector regression, instance-based learning, instance-based classifier, random forest, decision trees, and model trees for predicting geometallurgical properties from the grade of mining block. The predicted throughputs/hardness of materials from the above-mentioned methods can be used in the long-term planning models mentioned in Sect. 1.2.1. However, the predictions/calculations do not account for the blending of materials, which commonly happens in any mining operation.

Sepúlveda et al. (2018) proposed a method for material domaining that uses fuzzy clustering and graph-based labelling to generate clusters/domains in the mineral deposit that has similar geometallurgical properties. The method also accounts for the spatial location of the materials compared to k-mean clustering algorithms, which only accounts for material properties. Bhuiyan et al. (2019) proposed the use of the k-means clustering algorithm with random forest classifiers for modelling geometallurgical domains. The method uses k-means clustering algorithms on the first three principal components scores of isometric log-ratio coordinates of geochemical and geometallurgical (bond work index (BWI)) data to establish geometallurgical

domains. A random forest classifier is then used to identify the importance of geochemical data compared to geophysical data for geometallurgical domaining. Rajabinasab and Asghari (2019) studied the effect of k-means, hierarchical, and self-organizing maps clustering algorithms for generating clusters/domains of materials based on its geometallurgical properties. These geometallurgical domains generated from the above-mentioned methods can be used in the long-term planning models mentioned in Sect. 1.2.1 to schedule extraction of materials from specific geometallurgical domain of interest in a year. However, it's hard to determine which geometallurgical domain is of interest for a particular year.

Barnett and Deutsch (2012) discussed a systematic approach to use the transformation methods such as log ratios, MAF, normal score, stepwise conditional and introduced conditional standardization for transforming geometallurgical properties and then using such transformed values in Gaussian simulation methods to generate simulations of geometallurgical properties. Deutsch et al. (2016) presented an approach to handle the non-linearity, uneven sampling, and different scale measurements of geometallurgical properties. The proposed approach uses the Gaussian simulation method to handle non-linearity issues, the use of multiple regression of collocated properties within a super secondary approach (Babak and Deutsch, 2009) for modelling correlated continuous geological and geometallurgical properties at a different scales, uneven sampling is handled with a non-parametric Bayesian imputing method. Garrido et al. (2019) used a Gibbs sampler with rejection condition on non-linear geometallurgical recoveries at the processing mills to impute unevenly sampled geometallurgical properties. The method then uses Gaussian simulation methods conditioned on imputed geometallurgical samples to generate

simulations of geometallurgical properties. Ortiz et al. (2019) used the sequential Gaussian co-simulation method to generate simulations of geometallurgical property sag power index (SPI) and copper grades. The simulations of geometallurgical properties generated with the approaches mentioned above can then be used in the long-term planning models mentioned in Sect. 1.2.1. However, the long-term planning models cannot account for non-linearity with the blending of materials with different geometallurgical properties.

1.2.3   Short-term production planning

The long-term plan of a mining complex provides yearly plan and production targets, which are then used to develop a short-term plan. The short-term plan provides decisions about extraction sequence, destination policies, equipment assignment, and processing stream utilization at daily/weekly/monthly time scale and aims at maximizing the compliance to the yearly plan and production targets. Earlier work of Wilke and Reimer (1977), Fytas and Calder (1986), Mann and Wilke (1992), Schleifer (1996), Hustrulid et al. (2013) discusses some of the intricacies of short-term planning. A review of short-term planning can be found in Blom et al. (2018). The available literature can be categorized into three groups; first, models that determine the optimal allocation of fleets for a single mine operation, second, models that optimize the short-term extraction sequence and destination policies for a single mine operation, and third, models that simultaneously optimize fleet allocation, extraction sequence, and destination policies of a mining complex.

Li (1990) presented a deterministic LP model to determine the number of trucks required to haul the materials with minimum haulage cost, then a maximum inter

truck-time deviation rule is used to determine dispatching of trucks to a shovel. Alarie and Gamache (2002) discussed different solution strategies such as 1 Truck-N Shovel, M Trucks-1 Shovel, and M Trucks-N Shovels available for mine fleet allocation. Ta et al. (2005) presented a stochastic optimization model based on chance-constrained for the fleet assignment that accounts for uncertainty cycle time and truckload. Ercelebi and Bascetin (2009) used closed queuing network theory to determine the optimal number of trucks assigned to a shovel and then used a deterministic LP model to dispatch trucks to different shovels. Ta et al. (2013) proposed a stochastic MILP model for determining the truck allocation in an open-pit mine operation while accounting for shovel idle probabilities. Zhang and Xia (2015) presented a deterministic IP model that determines the optimal number and trips of trucks assigned to all shovels to minimize total truck operating costs and meet production targets. Nguyen and Bui (2015) used a multi-agent decision support system for the truck dispatching and solved it using a genetic algorithm. Kozan and Liu (2016) present a deterministic MILP model that determines timetabling, usage and scheduling of drilling, blasting, and excavating equipment to maximize the efficiency of mining equipment. Upadhyay and Askari-Nasab (2016) presented a deterministic mixed integer goal programming model that determines the allocation of shovel and trucks in an open-pit mine while maximizing shovel utilization and minimizing deviations from production targets and operating cost. Upadhyay and Askari-Nasab (2018) integrated a discrete event simulator in the mixed-integer goal programming model to determine shovel and truck allocation in an open-pit mine under uncertainty of the mining operation. Chaowasakoo et al. (2017) presented the use of global positioning system technology installed in trucks (Chaowasakoo et al., 2014) for

36

truck-shovel allocation in open-pit mines. Koryagin and Voronov (2017) presented a two-step approach that first determines the production target for each shovel material flow to different destinations and then uses a heuristic to determine the allocation of trucks to shovels that minimizes the total idle time. Ozdemir and Kumral (2019) used a simulation-based optimization method that first uses a discrete event simulator to determine an optimal number of trucks assigned to each mine and then uses a deterministic LP model to determine the allocation of each truck to different shovels in the pit. A review of models and algorithms used for mine fleet allocation and management can be found in Bielli et al. (2011) and Moradi Afrapoli and Askari-Nasab (2017).

Wilke and Reimer (1977) presented a deterministic LP model that determines the short-term extraction sequence for an iron ore mine and maximizes revenue subjected to blending and quantity constraints. Fytas and Calder (1986) proposed a deterministic LP model that determines the short-term extraction sequence to maximizes the revenue subjected to quantity, stripping ratio, and quality constraints. Fytas et al. (1993) presented a deterministic LP model that maximizes revenue considering head grade, concentrator tonnage, and stripping ration targets. Smith (1998) presents a mixed-integer goal programming model that optimizes the short-term extraction sequence to maximize ore production while minimizing the deviation from blending constraints. Kumral and Dowd (2002) presented a two-step deterministic model that uses integrated Langrangian parametrization and multi-objective simulation annealing framework to improve short-term extraction sequences generated from the Lerch and Grossman method (Lerchs and Grossman, 1965). The method also accounts for mining width accessibility constraints for equipment

operation. Gholamnejad (2008) presented a deterministic MILP model that optimizes the short-term extraction sequence while maximizing metal production and accounting for mining width accessibility, blending, slope, mining, and processing constraints. Huang et al. (2009) presented the MineSight Schedule Optimizer software that determines short and medium-term extraction sequences using a deterministic MILP model. Eivazy and Askari-Nasab (2012) proposed a deterministic MILP model that determines short-term extraction sequence and destination policies while minimizing total operating cost and account for blending, stockpile, horizontal mining direction, and ramp constraints.

L'Heureux et al. (2013) presented a deterministic MILP problem model that optimizes extraction sequence, displacement of shovels, drilling and blasting operations. The method uses batch constraints for scheduling drilling and blasting operations and minimizes the total operating cost. Mousavi et al. (2016a) present a deterministic MILP model that optimizes the extraction sequence and destination of mining blocks along with the assignment of the fleet. The model minimizes the total operating cost that includes re-handling, holding, misclassification and drop-cut costs. Drop cuts were defined as the blocks which are in the middle of surrounding blocks and were considered costly to extract and can be mined if new benches are required to be opened. The model also considered short-term stockpiles. A hybridized heuristic approach consisting of simulated annealing, large neighbourhood search and branch and bound is utilized to generate solutions with a deviation of 1 % from the solution achieved using exact methods for small instances problems. Mousavi et al. (2016b) further presented a comparative study of three different metaheuristics algorithms for solving the MILP model presented in Mousavi et al. (2016a). Three different

metaheuristics, namely, simulated annealing, Tabu search, and a hybrid Tabu search simulated annealing algorithm, were tested to solve the short-term production planning model. The hybrid algorithm performed better compared to the other metaheuristics. Blom et al. (2017) present a deterministic MILP model that determines the short-term extraction sequence, destination policies, and fleet assignment decisions. The model is solved using a multi-objective rolling horizon algorithm, which solves the model over the planning horizon sequentially. In addition, the solution approach first optimizes the short-term decisions to minimize deviation from the production target and then re-optimize such decisions to maximize equipment productivity. All the models mentioned above do not account for supply and equipment performance uncertainties. The two sources of uncertainty related to the supply of material and operation of different equipment affect the short-term operations. Matamoros and Dimitrakopoulos (2016) proposed a stochastic model that optimizes the extraction sequence and fleet assignment decisions while accounting for supply and equipment performance uncertainties, to meet the operational and blending requirements and minimize the cost of extracting and hauling materials, cost of equipment movement, lack of production from shovels, deviation from mineability and production targets. However, to reduce the computational cost aggregated continuous period, fixed mining direction and sequential solution approach was utilized. The models presented above for short-term production planning are limited to a single mine operation and do not optimize the several components of a mining complex simultaneously in one single optimization model while accounting for supply and equipment performance uncertainties.

Quigley and Dimitrakopoulos (2019) proposed a stochastic short-term planning model that optimizes the extraction sequence and fleet assignment decisions while accounting for supply and equipment performance uncertainties. The model minimizes the shovel movement cost, deviation from the different processing mill production targets, deviation from shovel production targets, and deviation from total available truck hours. Horizontal precedence relation that ensures a more practical and realistic extraction sequence is utilized in the model. Both and Dimitrakopoulos (2020) extend the models presented in Goodfellow and Dimitrakopoulos (2016) to a stochastic short-term planning model that optimizes the extraction sequence, destination policies, fleet assignment, and processing alternative decisions while accounting for supply and equipment performance uncertainties. The model maximizes cumulative cash flows while minimizing the deviations from processing mill production targets, shovel movements, shovel production targets, truck haulage hours, and mineability targets.

The models described above related to either short-term fleet allocation or production planning cannot learn from the incoming new information collected during the operations in a mining complex, and therefore cannot respond accordingly to adapt the short-term production plan.

1.2.4 Quantification of equipment performance uncertainty

Equipment performance uncertainty referred to the uncertainty associated with availability, utilization, productivity, break down, performance, repair time, cycle time, etc. associated with all the equipment used in a mining operation. For instance, the performance of shovels in an operating mine is uncertain in terms of load time,

productivity, availability, utilization, and more. Monte Carlo simulations are equally probable scenarios that are generated based on historical equipment performance data and can quantify the uncertainty associated with the performance of different equipment in a mining operation. These types of simulation methods do not require spatial statistics (a characteristic found in mineral deposit) but generate scenarios/simulations that reproduce the behaviour/statistics inferred from the historical data. There are two different types of approaches available in the literature to quantity equipment performance uncertainty. First, that estimates the parameters of an assumed distribution defined as a priori for equipment performance, and second, that generates empirical distribution without any assumption, using the historical equipment performance data.

Matamoros and Dimitrakopoulos (2016) generated simulations of shovel and truck availability, assuming a Gaussian distribution. Quigley and Dimitrakopoulos (2019) generated simulations of truck cycle time, also assuming a Gaussian distribution. Ozdemir and Kumral (2019) assume a PERT distribution to generate simulations of truck fill factor, loading time, dumping time, log-normal, normal and Weibull distributions for truck hauling time. Paduraru and Dimitrakopoulos (2019) generated simulations of the shovel load time, truck cycle time assuming a Gaussian distribution, shovel breakdown time assuming an exponential distribution, shovel repair time assuming a log-normal distribution. Both and Dimitrakopoulos (2020) generated simulations of shovel productivity and truck availability assuming a Gaussian distribution. Quigley and Dimitrakopoulos (2019) generate simulations of shovel production, utilization, availability, and truck utilization and availability with Monte Carlo simulation. The different correlated parameters (production, utilization,

and availability) of equipment performance are first decorrelated into independent factors using PCA (Hotelling, 1933); then, the PCA factors are simulated separately with Monte Carlo simulation methods; and finally, back-transformed into original space to generated simulated models of correlated equipment performance properties.

1.2.5   Incoming new information in mining complexes

The short-term plan is implemented to perform the day-to-day operations in a mining complex. New and conventional digital technologies, including advanced sensors and monitoring devices, allows a mining complex to monitor and collect new information in real-time from its different components during the day-to-day operation. The new information collected in a mining operation with conventional technologies includes mining equipment fuel consumption, engine power consumption, engine torque, engine temperature, production rates, efficiency, utilization, location and location of materials in a mining complex. The new information collected in a mining operation with advanced sensor include quality of materials extracted, transported, handled, conveyed, process and sold to the customers.

Koellner et al. (2004) discussed the use of conventional sensors installed on mining haul trucks that continuously monitors the status and performance of the trucks. These sensors can also monitor the fuel consumption, engine power consumption, brake efficiency, engine torque, engine temperature, engine speed, etc. of a truck. Conventional technology, such as the global positioning system (GPS), can pinpoint the location and status of the mining fleet in real-time (Chaowasakoo et al., 2014). The GPS data about the spatial location of the different equipment in a mining

operation can be used to improve the control systems (Sládková et al., 2011). Radio-frequency identification (RFID) tags are another example of conventional technology that can track and locate the flow of materials in a mining operation. Brewer et al. (1999) highlighted the scope to use such an RFID-GPS combined system to perform production scheduling in a job shop supply chain. Rosa et al. (2007) discussed the use of RFID tags during blasthole drilling to locate the displacement of materials during blasting and also the location of extracted material throughout the mining complex. Kargupta et al. (2010) discussed the MineFleet tool that monitors the health, fuel consumption, driver behaviours and gas emission from mining equipment. Wei et al. (2011) reviewed the critical conventional technology internet of things for labelling and precepting people, equipment, and environmental changes in an underground mine. Baek and Choi (2019) discussed the use of the conventional communication technology (ICT) system for tracking the location of equipment and people in real time and monitoring the environmental changes in an underground mine.

New digital technologies include IR (Larkin, 2011), and XRD (Jenkins and Snyder, 1996) sensors that generate surface measurements about mineralization of materials; XRF (Jenkins, 1999) and LIBS (Singh and Thakur, 2007) sensors that generate surface measurements about the grade and deleterious elements of materials; and DE-XRT (Eilbert, 2009) sensors that generate volumetric measurements about the grade, deleterious elements, and density of materials. Infrared (IR) sensors such as NIR/SWIR/MWIR passes a beam of infrared radiation through the surface of materials to excite their harmonic molecular vibrations. The molecules in the material absorb frequencies of the infrared beam that are characteristics of their mineralization. A detector then measures the amount of energy absorbed at each frequency. The

energy absorption profile at different frequencies is matched against a database of energy absorption profiles of known materials to find the type of mineralization of the materials. Some form of calibration techniques such as regression , PCA, partial least square, etc. is used to determine the composition of such mineralization.

XRD sensors bombard the surface of the material with X-ray beams and use a detector to measure the intensity and angle of diffraction of the X-ray beams. The detector and the X-ray beam are rotated through a range of angles to generate the intensity and diffraction profile. The intensity and diffraction profile are dependent on the mineralization of the materials. The profiles are matched against a database of diffraction profiles of known materials to find the mineralization and its concentration in the materials. XRF sensors bombard the surface of the materials with high energy X-rays, which causes the electron in the inner orbital of atoms to eject from their orbit. Electrons in the higher orbitals subsequently fall into the hole, causing to release energy in the form of photons. This photon energy depends on the type of atom and its concentration, which is measured by a spectrometer installed in the XRF sensors. XRF sensors, therefore, generates measurements about the concentration of different atoms by using this process. XRD and XRF data have higher precision compared to the IR, LIBS, and DE-XRT data because they require sample preparation and laboratory testing but are still considerably faster compared to geochemical laboratory analysis data. LIBS sensors consist of similar components, except that the material surface is ablated by a laser beam that breaks down the surface into a plasma consisting of atoms, ions, and free electrons. The spectrometer in LIBS sensors reads the radiation generated by electromagnetic radiations during the cooling of the plasma. The intensity of the radiation is related to the concentration of atoms and is

used to generate the surface measurements about the concentration of different atoms with the LIBS sensors. DE-XRT bombards broad-band radiation from an X-ray tube on the materials. The X-ray radiations penetrate the materials and get absorbed. The degree of absorption depends on the material composition, specifically its atomic density and thickness. The intensities of the transmitted attenuated X-ray is detected by two sensors that capture different energy level allowing to characterize the material by its atomic density. The scanners generate a grayscale image showing the fraction of high (dark) and low (light) atomic density.

Goetz et al. (2009) used the NIR to detect alteration (swelling clays, kaolinite, muscovite, and biotite) of materials transported through conveyor belts in a mining operation. The NIR measurements were then used with XRD and XRF data in a regression model to predict the concentration of such alterations. Dalm et al. (2014) used the NIR sensors to determine the alternation (white mica, chlorite, tourmaline, and kaolinite) of drillhole core samples at a copper porphyry deposit. The results from NIR sensors showed good validation results against XRD sensor data. A multivariate logistic regression model is then used to determine the relationship between copper grade data acquired using XRF sensors and the alteration data acquired from NIR sensors. Iyakwari et al. (2016) used the NIR sensor to detect the alternation (chrysocolla, muscovite, kaolinite, biotite, chlorite, tourmaline, hematite, malachite, calcite, ankerite, and apatite) of run-of-mine samples collected from a copper deposit. The NIR sensor data showed good validation results against the XRD sensor data. The detected alternation from NIR data is used to classify materials as waste, ore, and middling and showed a good validation against the classification of materials with cut-off grades using XRF sensor grade data. Dalm et al. (2017) used the SWIR sensors

to determine the alteration (white mica, chlorite, tourmaline, nontronite, and kaolinite) of drillhole core samples at a copper deposit. Subsequently, a PCA method is used to find and analyze the correlation between grade data from XRF sensors and alternation data from SWIR sensors. Dalm et al. (2018) used the NIR and SWIR sensors to determine the alternation (iron oxide, pyrite, pyrophyllite, alunite, dickite, illite, zunyite, jarosite, water, and diaspore) of drillhole core samples at an epithermal gold-silver deposit. The results of the NIR and SWIR detections showed a good validation against the XRD data of the drillhole core samples. A partial least square discriminant analysis method is then used to associate NIR and SWIR data to geochemical laboratory analysis data.

Cremers (1987) proposed and used the LIBS sensor to determine the grade (manganese, chromium, silicon, copper, and iron) of steel products. Bolger (2000) used the LIBS sensor for determining the grade (arsenic, copper, chromium, cobalt, iron, manganese, nickel, and zinc) of drillhole core samples. A normalization scheme based on total plasma emission is used to correct variations in radiation measurements of uneven drillhole core samples. The LIBS sensor data showed a good correlation with the geochemical laboratory analysis data. Bette et al. (2005) used the LIBS sensor to measure the grade of steel products. Elements such as carbon, nitrogen, oxygen, phosphorus, and sulphur are detected in the steel products, which leads to the identification of non-metallic inclusion about aluminum nitride, aluminum oxide in steel products. Death et al. (2008, 2009) used the LIBS sensor to determine grade (iron, aluminum, silicon, manganese, potassium, phosphorus) of run-of-mine samples from iron ore deposits and grade (aluminum, cobalt, chromium, iron, manganese, magnesium, nickel, chlorine, sodium, sulphur, silicon, and titanium) of drillhole core

46

samples at a nickel laterite deposit. A principal component regression (PCR) model is used to build a relationship between LIBS sensor data and XRF sensors grade data. The PCR model showed a good correlation with the XRF sensor grade data.

De Jong (2004) used the DE-XRT sensors to determine ash content, zinc grade, and lead grade of samples from the conveyor belt. The authors suggested using such sensors for the sorting of materials fed to the processing mill. Lessard et al. (2014) used the DE-XRT sensor to determine the molybdenum grade of run-of-mine samples at two molybdenum deposits. The authors further showed that using DE-XRT data for sorting the materials fed to the processing mill can help to reduce the operating cost by 30% by processing less materials with high molybdenum grades, increases the life of mine by identifying profitable materials during sorting procedure, and determine operating expenditure requirements to further increase the throughput and recovery of molybdenum materials. Kern et al. (2019) used the DE-XRT sensor to determine the abundance (grade) of mineral cassiterite (primarily contains tin), and the SWIR sensor to determine the alternation (chlorite) in run-of-mine samples at a tin-indium-zinc deposit. The DE-XRT and SWIR data were used to build proxies for the mineral liberation analysis data (mineralogy, mineral density, mineral area, mineral grain size, etc.), which is commonly used for sorting materials fed to the processing mill.

The new information collected either with the conventional technologies and advanced sensors can allow a mining complex to learn and respond accordingly to update the supply and equipment performance uncertainties and subsequently adapt the short-term production plan. However, the technologies outlined above only allows for the collection of the new information, but do not use the incoming new information

to learn and respond accordingly to update supply uncertainty and adapt short-term production planning decisions.

1.2.6 Updating supply and equipment performance uncertainties with incoming new information

The new information collected from the different components of a mining operation can be transmitted using Wi-fi networks (Boulter and Hall, 2015) to be stored in a central system containing internet database management platforms (Ramírez-Gallego et al., 2018). The stored data can then be retrieved, whenever required, to update the supply and equipment performance uncertainties. However, methods to update the supply and equipment performance uncertainties should account for the "soft" characteristic of the incoming new information. This is a well-known issue in petroleum and groundwater reservoirs (Oliver et al., 2008), also known as history matching, where uncertain production data (oil production, flow rates, well pressure, seismic data, tracer observations, etc.) is used to update the set (simulations) of static (such as porosity, permeability, etc.) and dynamic reservoir variables (such as pressure, fluid saturation, etc.), such that the updated variables better matches the observed behaviour (production data). Oliver and Chen (2011) provided a review of the methods used for history matching.

Ensemble Kalman filter (Evensen, 2009) is a well-known, studied, and applied method for history matching in groundwater and petroleum reservoirs (Aanonsen et al., 2009; Oliver and Chen, 2011). Kalman filter (Kalman, 1960; Kalman and Bucy, 1961) was used initially for updating reservoir models with new information. However, the method was limited to linear models and was improved in Jazwinski

(1970) and Brown and Hwang (1992), to introduce extended Kalman filter (EKF) for small-scale nonlinear models. Evensen (1994a) introduced the ensemble Kalman filter (EnKF) method and used it (Evensen, 1994b) to assimilate information in large-scale nonlinear oceanic models. The performance of EnKF and EKF is compared in Miller et al. (1999) and Madsen and Canizares (1999). Other applications of EnKF include water model (Van Loon et al., 2000), stock assessment model (Grønnevik and Evensen, 2001), atmospheric circulation model (Houtekamer and Mitchell, 2001; Mitchell et al., 2002).

$$y_{t,j}^p = f\left(y_{t-1,j}^u\right) \qquad , \forall j \in [1, N] \quad (1.11)$$

$$d_{obs,t,j} = d_{true,t} + \epsilon_t \qquad , \forall j \in [1, N] \quad (1.12)$$

$$y_{t,j}^u = y_{t,j}^p + K_t\left(d_{obs,t,j} - H_t y_{t,j}^p\right) \qquad , \forall j \in [1, N] \quad (1.13)$$

$$K_t = P_t^p H_t^T (H_t P_t^p H_t^T + C_{D,t})^{-1} \qquad (1.14)$$

$$P_t^p = \frac{1}{N-1} \sum_{j=1}^{N} \left(y_{t,j}^p - \overline{y_{t,j}^p}\right) \cdot \left(y_{t,j}^p - \overline{y_{t,j}^p}\right)^T \qquad , \forall j \in [1, N] \quad (1.15)$$

EnKF consists of two main steps, the forecast step and the assimilation step. The forecast step represented by Eq. 1.11 generates the predictions at time step t about the state vector based on all the information available prior to time step t. The updated state vector, $y_{t-1,j}^u$, at time step $t-1$ for simulation number j contains variables required to describe the system, i.e. model variables (static and dynamic) and predicted observed data. f represent the reservoir flow simulator, p denotes predicted, u denotes updated, j is the simulation index, N denotes the total number of simulations, t is the time step index, and $y_{t,j}^p$ is the predicted state vector. The forecast

step only changes the dynamic variables and the predicted observed data, but the static variables remain unchanged in the forecast step. The observed ($true$) data, $d_{true,t}$, measured with sensors at time step t is perturbed by adding noise to reflect the uncertainty in the measurement with sensors, using Eq. 1.12. $d_{obs,t,j}$ denotes the perturbed observation (obs) at time step t and for simulation number j. ϵ_t represent the measurement noise and is usually assumed to be Gaussian. The second step in the EnKF method is the assimilation step that adjusts all the variables in the state vector to honour the observed data using Eq. 1.13. H_t is an operator that extracts the components of the state vector corresponding to the observed data. K_t is the Kalman gain matrix that weighs the relevance of observed data compared to the state vector and is computed using Eq. 1.14. $C_{D,t}$ is the covariance matrix of measurement error, $C_{D,t} = E[\epsilon_t \cdot \epsilon_t^T]$, and is typically assumed to be diagonal. P_t^p is the covariance matrix of the predicted state vector, computed using Eq. 1.15, and consist of covariance of model variables, the covariance of predicted observed data, and cross-covariance between model variables and predicted observed data.

Variant of EnKF such as EnKF with ensemble Kalman Smoother (Skjervheim et al., 2007), EnKF with iterative confirming option (Wen and Chen, 2005), EnKF with confirming option that update the dynamic variables based on updated static variables to achieve consistent dynamic and static parameter (Wen and Chen, 2006), EnKF with Gaussian mixture model (Dovera and Della Rossa, 2011), EnKF with iterative ensemble smoother (Chen and Oliver, 2012), EnKF with multi-point simulation method (Hu et al., 2013), combination of EnKF with pilot point and

gradual deformation parameterization technique (Heidari et al., 2013), gathered EnKF (Shuai et al., 2016), were also utilized for updating different nonlinear models.

Other methods for history matching of petroleum and groundwater reservoir includes gradual deformation (Hu, 2000), neighbourhood methods (Sambridge, 1999), evolutionary algorithms (Schulze-Riegert and Ghedan, 2007), gradient-based methods (He et al., 1997), maximum a posteriori (Oliver, 1996), Markov chain Monte Carlo (Oliver et al., 1997), randomized maximum likelihood (Sarma et al., 2006; Chen and Oliver, 2012; Vo and Durlofsky, 2014; Shirangi, 2017), Tau-model (Naraghi and Srinivasan, 2015), multi-point statistics and convolution neural network (CNN) with PCA (Liu et al., 2019), stepwise CNN-PCA and recurrent neural network (RNN) approach (Tang et al., 2019), and Markov mesh models (Panzeri et al., 2016). The CNN-PCA method is highly subjective to training images used in the method to train the CNN and will perform adversely when the static variables (captured in the training images) of the reservoirs changes significantly. Similarly, the CNN-PCA-RNN method uses the results of the flow simulator over the initial simulations, which makes the RNN model highly subjective to initial simulations. This method will perform adversely when the initial simulations and, subsequently, the results from flow simulation change.

Unlike petroleum and groundwater reservoirs, mineral deposits only consist of static variables (geological properties) that needs to be updated with incoming new information. Benndorf (2015) introduced the use of EnKF for updating estimated orebody models. Benndorf and Buxton (2016) presented a closed-loop conceptual framework for real-time reserve management that integrates the sensor-based measurement to update the supply uncertainty of mineral deposits. The framework is

tested at a Walker Lake dataset (Isaaks and Srivastava, 1989) to show the efficiency of EnKF compared to Kalman filter for one, two, and three shovels simultaneous operations in a mine. Yüksel et al. (2016) used the EnKF method to update supply uncertainty (ash content) with XRD ash content data of drillhole samples at a coal deposit. The method also incorporates a rejection sampling technique to validate and reject some of the updated simulations. Yüksel et al. (2017) used the EnKF method to update the supply uncertainty (ash content) from radiometric measuring system sensor data (ash content) of materials passing on conveyor belt at a coal deposit. The authors further showed that initial simulations generated by either adding noise to estimated model or using SGS are updated with EnKF with similar performance.

Wambeke and Benndorf (2017) used the EnKF method with a forward simulator (similar to reservoir simulator) to update the supply uncertainty with sensor data of conveyor belt materials at a synthetic mining operation. The method also incorporates a connected updating cycle technique to split the updating equations of EnKF and a local neighbourhood while updating to reduce the computational requirements of the EnKF covariance calculations. Wambeke and Benndorf (2018) studied the effect of measurement volumes, blending ratio, and sensor precision within the EnKF method to update supply uncertainty at a synthetic mining operation. The study showed that optimal conditions for using the EnKF method for maximizing the potential of updating supply uncertainty are when the measurement of sensor data are recorded on small extracted volumes of materials from single shovel with high sensor precision and minimal blending. Wambeke et al. (2018) used the EnKF method to update the supply uncertainty (geometallurgical property BWI) at Tropicana gold mine with the processing mill conventional sensor data (throughput, power draw, feed

and product size). The method used a forward simulator to mimic the mining operation. The simulator provides the estimates about the characteristics of the materials fed to the mill and then the Bond's work theorem of comminution (Bond, 1952) is used to estimate the power and throughput of the mill. The EnKF method updates the supply uncertainty (SGS simulations of BWI property conditioned on XRF and NIR drillhole samples) to update the historical and future forecasts by 72% and 26%, respectively. Other methods for updating supply uncertainty of mineral deposits with new information include conditional simulation by successive residual (Vargas-Guzmán and Dimitrakopoulos, 2002, 2003; Jewbali, 2006; Jewbali and Dimitrakopoulos, 2011) and co-simulation with soft data (Dimitrakopoulos and Kaklis, 2001; Neves et al., 2018). The co-simulation methods are based on developments in petroleum reservoirs to simulate static variables with soft seismic data (Journel and Alabert, 1990; Zhu, 1991; Verly, 1993; Fichtl et al., 1997; Mao and Journel, 1999b; Soares, 2001; Horta and Soares, 2010; Tahmasebi and Sahimi, 2015; Soares et al., 2017).

Equipment performance uncertainty related to its utilization, availability, downtime, repair time, productivity etc. can be updated with the new information collected from the equipment during the mining operation. The available literature regarding updating equipment performance uncertainty is limited. Upadhyay and Askari-Nasab (2018) and Ozdemir and Kumral (2019) suggested using the equipment production sensor data to update the coefficients of the distributions that quantify the equipment performance uncertainty.

The EnKF method outlined above to update supply uncertainty with incoming new information is limited to (a) updating of Gaussian variables, therefore imposes

Gaussian assumptions to the geological properties of mineral deposits, (b) cannot learn from the incoming new information, and, (c) cannot account for high-order spatial statistics in the incoming new spatial information. Updating equipment performance uncertainty with methods outlined above cannot integrate the incoming new information that does not fit the distribution that was used to quantify the equipment performance uncertainty.

1.2.7 Responding to incoming new information using machine learning in industrial complexes

The incoming new information, along with the updated supply and equipment performance uncertainties, can be used to make real-time decisions in a mining complex. Traditionally, the blasthole drilling data is used to define destination of materials using cut-off grade policies (Verly, 2005; Abzalov et al., 2010; Dimitrakopoulos and Godoy, 2014; Vasylchuk and Deutsch, 2018), equipment GPS data is used to decide for dispatching decisions (Kargupta et al., 2010; Nguyen and Bui, 2015), and trend analysis of equipment performance to determine the appropriate time for maintenance (Myrzabekova et al., 2020). However, all the above-mentioned methods do not (i) integrate all the incoming new information, (ii) account for the "softness" in the incoming new information, (iii) learn from the incoming new information, (iv) account for interactions of components in a mining complex, and (v) make all the short-term decisions simultaneously in a mining complex. The combinatorial optimization methods, although they can simultaneously optimize all the short-term decisions but cannot use the incoming new information to learn and respond accordingly to make such short-term decisions. Machine learning algorithms

are known for their ability to learn how to make decisions based on incoming new information. These algorithms are classified into three different categories, supervised, unsupervised, and reinforcement learning. Supervised learning is learning from a training set of labelled examples (Hastie et al., 2009). Each example in the training set consists of predictors (features that describe the situation) and label (the correct action). The objective of supervised learning is then to learn to extrapolate or generalize the responses so that it acts correctly in unseen situations. Unsupervised learning tries to find hidden patterns and structures in the collection of an unlabeled training set that only contains predictors (Hastie et al., 2009).

Supervised machine learning algorithms have been used in mining for determining rock types (Chatterjee, 2013; Patel and Chatterjee, 2016), reliability estimation of mining equipment (Chatterjee et al., 2015), prediction of equipment failure (Campeau and Dubois, 2019), prediction of fuel consumption by trucks (Siami-Irdemoosa and Dindarloo, 2015), detecting objects in shovel bucket (Shariati et al., 2019), predicting fly rock (Amini et al., 2012; Khandelwal and Monjezi, 2013a), back break (Khandelwal and Monjezi, 2013b), and ground vibration (Iphar et al., 2008; Monjezi et al., 2013) during blasting operation, predicting landscape changes (Demirel et al., 2011; Mountrakis et al., 2011; Petropoulos et al., 2013), performance of hydro cyclones (Karimi et al., 2010), jigging (Panda et al., 2012), and vibrating table concentrator (Panda and Tripathy, 2014) in processing mills, and predicting groundwater inflow in open pit mines (Bahrami et al., 2016). Unsupervised machine learning algorithms have been used in mining for defining clusters/domains of materials in mineral deposits based on their geometallurgical properties (Sepúlveda et

al., 2018; Lishchuk et al., 2019; Bhuiyan et al., 2019; Rajabinasab and Asghari, 2019) and grade properties (Goodfellow and Dimitrakopoulos, 2016, 2017).

Unlike supervised and unsupervised machine learning, reinforcement learning (RL) algorithms aim at generating agents that learn to make decisions given some observations, through its own experience by interacting with an environment that provides feedback, to maximize/minimize the long-and-short-term goal/reward/objective/value. An agent is usually a function approximator, defined using parameters θ, such as neural networks, deep neural networks, convolution neural networks, decision trees, and so on. An environment is a model of the industrial process that encapsulates the critical intricacies of its operations and is responsible for providing feedback to the actions taken by the agent. An industrial process is usually assumed to be a Markov decision process (MDP). An MDP is a discrete time-step stochastic process, where at each time-step, a decision-maker/agent makes an action based on the state provided by the environment and one time-step later receives a scalar reward and next state from the environment. The probability that the process transition to a new state is only dependent on immediately preceding state and action but is independent of the rest of the previous states and actions.

Let's assume an environment represents a mining complex, and the agent is denoted by f_θ, where θ are the parameters/policy of the agent. For example, the environment of a mining complex will encapsulate the intricacies of its day-to-day operation. The agent receives an observation S_t (states) at time t, from the environment and usually assumed to be a fully observable Markov decision process (Figure 1-3). For example, states in a mining complex can be quantity and quality of

56

materials available, handled, and processed; quantity and performance of equipment available and used.

Figure 1-3 The agent-environment interaction in an MDP (Sutton and Barto, 2017)

The agent receives the state S_t from the environment and the takes/proposes an action $A_t = f_\theta(s_t)$. For example, actions in a mining complex can be the short-term decisions like whether to extract the materials, what equipment to use to extract the materials, where and how to process the extracted materials, etc. The action taken/proposed by the agent is then executed by the environment to transition to the new state S_{t+1}. The environment also provides a scaler reward R_t as a feedback for actions taken/proposed by the agent. For example, in a mining complex, this can be profit/cashflows/costs/metal production/equipment queue. The RL algorithm then teaches the agent to find a policy, given the data generated by the interaction of the agent with the environment, to maximize the expected future reward. Alternatively, some RL algorithms try to find an agent that maximizes the reward of the state-action pair and then takes the action which has the highest state-action reward.

Deep reinforcement learning (DRL) uses a deep neural network as an agent in the RL algorithm. DRL algorithms are classified into two categories, first, model-based

DRL that has access to the model of the environment and second, model-free DRL that do not have access to the environment. Since model-based DRL has access to the environment model, they can plan better by looking ahead in the future, usually done via Monte Carlo rollouts that perform simulations to see what happens in the future. Model-free DRL, on the contrary, are usually sample inefficient (transitions from one state to another are not sampled efficiently by the agent) but are usually scalable and easier to implement and tune for different types of planning tasks. Ivanov and D'yakonov (2019) and Arulkumaran et al. (2017) provided a survey about the different types of DRL algorithms.

Paduraru and Dimitrakopoulos (2018) proposed a model-free Bayesian RL algorithm to optimize the destination policies of materials in a mining complex. However, the method developed requires a predefined extraction sequence to calculate the expected a posteriori improvement in the objective function during the optimization. Paduraru and Dimitrakopoulos (2019) proposed a model-free policy gradient RL algorithm to optimize the neural network destination policies of materials in a mining complex while accounting for supply and equipment performance uncertainties. The neural network destination policies increased the expected cash flows by 6.5 % compared to the mine's cut-off grade destination policies for a copper mining complex. However, the method is (a) limited to a single product mining complex, and (b) does not provide a required continuous updating of the short-term production plan regarding destination policies of materials with the incoming new information generated from sensors.

Application of model-free DRL algorithms in different areas of planning include scheduling tasks in manufacturing and industrial environments (Aissani et al., 2012;

Leal et al., 2019), scheduling tasks and jobs to computer resources (Mao et al., 2016), detecting and defending against cyber attacks (Li et al., 2019), managing financial portfolios (Sato, 2019), controlling traffic lights to avoiding congestion (Arel et al., 2010), controlling network access in communication networks (Luong et al., 2019), controlling movement of robotic arm (Kober et al., 2013; Lillicrap et al., 2015; Levine et al., 2016), maintenance of truck fleet (Barde et al., 2019), ads bidding and placement (Jin et al., 2018; Zhao et al., 2019), news recommendation (Zheng et al., 2018), web system configuration (Bu et al., 2009), playing games (Mnih et al., 2013; Guo et al., 2014; Hausknecht and Stone, 2015; Lillicrap et al., 2015; Baker et al., 2019), learning to paint like humans (Ganin et al., 2019), etc. Model-based DRL algorithms are used for playing games (Silver et al., 2016, 2017, 2018; Schrittwieser et al., 2019), finding experimental conditions for chemical reaction trials (Zhou et al., 2017), learning to paint like humans (Huang et al., 2019), etc.

The above-mentioned RL algorithms are suitable for learning how to make short-term production planning decisions in a mining complex to respond to the incoming new information since a labelled training data set is not available. However, RL algorithms have not been used to develop a self-learning mining complex, where agents learn to respond to the incoming new information to update the supply uncertainty of mineral deposits and then adapt the short-term production plan of a mining complex.

1.3   Goal and objectives

The goal of this thesis is to develop a self-learning mining complex that can update its supply and equipment performance uncertainties and then adapt its short-term production plan in real-time by responding to the incoming new information.

The following objectives are addressed to achieve this goal:

(1) Review and outline limitations of past work related to stochastic integer programming, simultaneous optimization, modelling and updating of supply and equipment performance uncertainties, and reinforcement learning algorithms in mining and other areas of planning.

(2) Expand the simultaneous stochastic optimization model for long-term production planning in mining complexes to account for supply uncertainty related to continuous non-additive geometallurgical properties (SPI and BWI) of materials.

(3) Develop a self-learning mining complex that uses incoming new spatial information about properties of materials extracted to update the supply uncertainty with an extended EnKF, and learn to adapt the short-term destination of extracted materials with an agent trained via a model-free policy gradient reinforcement learning algorithm, within the long-term production plan developed in objective (2).

(4) Extend the self-learning mining complex in objective (3) to use the incoming new information about equipment performance to update the equipment performance uncertainty with a Monte Carlo simulation algorithm. Subsequently, develop a new model-based self-play deep reinforcement learning algorithm to train agents to learn to adapt the short-term production plan of a mining complex (extraction sequence, destination of extracted

materials, and processing stream utilization decisions simultaneously) within the long-term plan developed in objective (2).

(5) Develop a new model-free deep reinforcement learning algorithm replacing the extended EnKF in objective (3) to train agents that learn to update the supply uncertainty of materials with incoming new spatial and temporal information while accounting the high-order spatial statistics of available data.

(6) Outline the contributions and limitations of the developed methods and provide directions for future work.

1.4   Thesis outline

This thesis is organized into the following chapters:

Chapter 1 provides a general overview and motivation behind this thesis. It also provides the relevant literature review related to all the topics covered in this thesis, which include simultaneous stochastic optimization of long-term and short-term production plan of mining complexes, quantification of supply and equipment performance uncertainties, incoming new information collected in a mining complex, methods to update supply and equipment performance uncertainties with incoming new information, and reinforcement learning algorithms for making decisions by responding to the incoming new information. This chapter also states the goal and objectives of the thesis.

Chapter 2 describes an approach to integrate supply uncertainty related to two continuous non-additive geometallurgical properties (SPI and BWI) of materials in the simultaneous stochastic optimization of the long-term production plan of a mining complex. The results of the proposed method at a copper mining complex are

compared with a long-term production plan optimized using industry-standard practices to show its benefits and added value.

Chapter 3 develops a self-learning mining complex for adapting the destination of materials in a mining complex with incoming new information within the long-term plan developed in Chapter 2. The framework first uses an extended ensemble Kalman filter (EnKF) that combines minimum/maximum autocorrelation factors with EnKF to update the supply uncertainty about multiple properties of materials with incoming new spatial information. The updated supply uncertainty, along with equipment performance uncertainty, is then fed to a neural network agent trained via a model-free policy gradient reinforcement learning algorithm to adapt the short-term flow of materials in a mining complex. An application of the proposed framework at a copper mining complex with multiple mines, destinations, processing streams, and products demonstrates the benefits of the proposed framework.

Chapter 4 develops a model-based self-play reinforcement learning algorithm for adapting all the relevant short-term production planning decisions simultaneously (extraction sequence, the destination of materials, and processing stream utilization) in a mining complex with incoming new information within the long-term plan developed in Chapter 2. The self-play reinforcement learning algorithm combines a Monte Carlo tree search with a deep neural network agent. The proposed algorithm is integrated into the self-learning mining complex proposed in Chapter 3. The new self-learning framework uses the extended EnKF to update the supply uncertainty and a proposed Monte Carlo simulation method to update the equipment performance uncertainty. The updated supply and equipment performance uncertainties is then fed to the deep neural network agent trained via the proposed self-play reinforcement

learning algorithm to adapt all the short-term planning decisions simultaneously. The new framework is applied at a copper mining complex with multiple mines, destinations, processing streams, and products to demonstrates its efficiency, benefits and applied aspects.

Chapter 5 develops a model-free deep deterministic policy gradient reinforcement learning algorithm replacing the extended EnKF in Chapter 3 to train agents that learn how to update the supply uncertainty of materials with incoming new spatial and temporal information about properties of materials. The proposed algorithm combines high-order spatial statistics with an actor and critic convolutional neural network agents within a deep deterministic policy gradient reinforcement learning algorithm. An application of the proposed algorithm at a copper mining operation with incoming spatial sensor data from drilling machines, and temporal processing plant sensor data, demonstrates its applied aspects in updating the supply uncertainty while reproducing spatial patterns and high-order spatial statistics.

Chapter 6 outlines the contributions, summarizes the findings of the thesis, and provides recommendations for future work.

CHAPTER 2
Application of Simultaneous Stochastic Optimization with Geometallurgical Decisions at a Copper-Gold Mining Complex

This chapter introduces an approach to integrate uncertainty and decisions about two non-additive geometallurgical properties (semi-autogenous power index (SPI) and bond work index (BWI)) in the simultaneous stochastic optimization of the long-term production plan of a mining complex.

2.1 Introduction

A mining complex consists of multiple components such as mines, crushers, stockpiles, leach pads, processing mills, waste dumps, means of transportation and customers. The components are interlinked to form the mineral value chain. Material extracted from the mines flows through the mineral value chain to finally produce the products which are delivered to customers and/or spot markets. In the last decade, developments have been made to integrate different decisions in the mineral value chain into one single optimization model. The integrated model of a mining complex simultaneously optimizes the different decisions in the related mineral value chain to utilize synergies by jointly developing the sequence of extraction, destination policies, processing stream utilization, operating modes, transportation alternatives, capital expenditure, and so on.

Past work in simultaneous optimization models of mining complexes can be categorized into two groups: conventional approaches and stochastic approaches.

Hoerger et al. (1999) present a model for Newmont's Nevada mining complex that simultaneously optimizes the timing of open-pit layback, underground stope development, capital expenditure, the timing of processing plant startup and shutdown, and, material routing decisions. Stone et al. (2007) present BHP's advanced mine planning optimization tool, "Blasor," which simultaneously optimizes the long-term panel extraction sequence and amount of material extracted from multiple mines. Whittle and Whittle (2007) present a global asset optimization model that includes optimization of the extraction sequence, mining rate, cut-off grade policy, processing path selection, and stockpiling strategy to satisfy production targets. The model does not optimize the different components simultaneously. It instead repeatedly creates random feasible extraction sequences and finds a locally optimal blending and processing stream utilization strategy. Whittle (2010, 2014) presents the ProberC algorithm that generalizes the algorithm used to solve the global asset optimization model. In ProberC, grades are expressed as an additive quantity, such as metal quantity. This structure allows the integration of aspects of stockpiling, transportation, and so on, in the global asset optimization model. An integrated approach is presented by Pimentel et al. (2010) to address the simultaneous optimization of mining complexes and possible solution strategies. Some of the typical limitations of the conventional approaches include: (i) the aggregation of mining blocks into larger volumes before starting the optimization process to reduce computation requirements, (ii) not simultaneously optimizing the different components of mining complex, i.e. having separate models for different components of mining complexes that interact through some form of heuristic approach, (iii) not

65

considering uncertainties in grade, geometallurgy and material types of the mineral deposits.

Simultaneous stochastic optimization of mining complexes overcomes the abovementioned limitations of earlier models by considering one single optimization model to simultaneously optimize the different components of a mining complex under uncertainty. Montiel and Dimitrakopoulos (2015) propose a model that simultaneously optimizes block extraction, block destination, processing stream utilization, operating modes, and transportation alternatives, under grade and material type uncertainties. Results from a copper-gold mining complex indicate deviations of less than 3% and 1.2% from capacity and blending targets respectively for the proposed model compared to large and impractical deviations in the range of 30 to 40 % and 11 to 22 %, respectively, in the conventional mine production plan (Gemcom, 2005). In addition, the proposed method increases the net present value (NPV) by 5% as compared to the conventional mine production plan. Montiel et al. (2016) extend the model to include underground mines. Montiel and Dimitrakopoulos (2018) present another extension integrating the specification of additional practical operational constraints to mine production schedules, as well as a large-scale application and comparisons with conventional methods at the Twin Creeks gold mining complex, Nevada. Montiel and Dimitrakopoulos (2017) present a metaheuristic algorithm to solve the large optimization model of mining complexes. The algorithm first changes the block extraction and destination decisions to improve the NPV and then selects the optimal operating and transportation modes. The decisions are synchronized as the algorithm progresses. The proposed method increases the NPV by 30% and decreases the deviation from production targets by

66

45% compared to the conventional mine production plan. Goodfellow and Dimitrakopoulos (2016) present a model that simultaneously optimizes block extraction decisions, multiple element destination policies, and processing stream utilization strategies, under grade and material type uncertainties. The proposed method is applied at a copper-gold mining complex and shows deviations of less than 10% compared to 40% in the conventional mine production plan while increasing the NPV by 22.6%. Goodfellow and Dimitrakopoulos (2017) study the efficiency of the multiple element destination policies. Results from a nickel-laterite mining complex show a deviation of less than 1% for blending targets and an improved NPV of 3% for multiple element destination policies under supply uncertainty compared to multiple element destination policies for estimated deposits. Some of the common and useful aspects of simultaneous stochastic optimization models of mining complexes can be summarized as: (i) the ability to capitalize on synergies between different components of mining complexes, (ii) the capability to take advantage of and capitalize on the variability of grade and material type, (iii) not aggregating mining blocks into large volumes to misrepresent mining selectivity, (iv) focusing on the value of product sold to the different customers instead of the value of mining blocks, (v) increased expectation of meeting production forecasts, and (vi) higher NPV and metal production.

The stochastic mine planning models reviewed above, although they do incorporate grade and material type uncertainties, do not integrate the uncertainty and variability related to geometallurgical properties of the material. Geometallurgical properties such as grinding ability/hardness of material processed, ore texture, and so on, affect the throughput and energy consumption of the comminution circuits. Not

accounting for such geometallurgical aspects in strategic mine planning models may result in sub-optimal mine plans (Dowd et al., 2016). More specifically, comminution circuits in a mining complex typically account for 33% to 50% of the mine costs (Curry et al., 2014), which depend on factors such as the feed particle size, hardness of the material and many more. Ballantyne and Powell (2014) present a survey of the comminution circuit energy consumption of a major Australian gold and copper producer which accounts for 1.3% of Australian electricity consumption. The hardness of the material is measured using grindability tests such as the semi-autogenous (SAG) power index (SPI) for SAG mills and the Bond work index (BWI) for ball mills. Verret et al. (2011) present the overview of the different grinding ability tests used, characterizing ore hardness for comminution circuits. Hardness indexes such as the SPI and BWI are hard to model due to their sparse sampling nature. Different approaches are proposed to model geometallurgical properties such as a primary and secondary response framework (Coward et al., 2009), a combination of simulation and power approaches (Alruiz et al., 2009), the use of two-stage linear regression models (Boisvert et al., 2013), a projection pursuit method (Sepulveda et al., 2017), and fuzzy clustering (Sepúlveda et al., 2018). However, it is even harder to include the hardness indexes such as the SPI and BWI in the strategic mine planning models due to their non-additive nature. In addition, blending of material with substantially different hardness properties can introduce the problem of inconsistent throughput and higher energy consumption in the comminution circuits.

The work presented herein introduces an approach to integrate the geometallurgical uncertainty of two hardness indexes, the SAG power index (SPI) and the Bond work index (BWI), into the simultaneous stochastic optimization model

of mining complexes. In addition, the proposed model also deals with the blending of multiple materials with different hardness properties to ensure a consistent throughput of material to the comminution circuits. In the following sections, a brief description of the extended model of Goodfellow and Dimitrakopoulos (2016) is provided, followed by an application at a large copper-gold mining complex that consists of 2 mines with 12 different material types, 5 crushers, 2 stockpiles, 3 processing mills, 1 waste dump, and 2 leach pads, and produces copper, gold, silver, and molybdenum products. Conclusions and future work follow.

2.2   Method

This section discusses the model for simultaneous optimization of mining complexes under uncertainty proposed in Goodfellow and Dimitrakopoulos (2016), which is extended in this work to include uncertainty and constraints related to geometallurgical properties associated with grindability (SPI and BWI). Geometallurgical properties are non-additive in nature and this specific aspect is not accounted for in Goodfellow and Dimitrakopoulos (2016). The work presented herein provides an approach to include constraints and uncertainty related to such non-additive geometallurgical properties in the context of the simultaneous optimization of mining complexes. In addition, and most importantly, the work presented highlights the applied aspects of the method with an application at a large copper-gold mining complex with over 1.6 million binary and 50,000 continuous variables. First, definitions and notation are provided, then, the different decision variables are discussed. Finally, the objective function and constraints are outlined.

2.2.1 Definitions and notations

The notations used in this section are as follows. $\mathbb{M}$ represents a group of mines. $\mathbb{B}_m$ represents the set of mining blocks for mine m and $MC_{b,t}$ represents the mining cost of a block b in period t. A block b is accessible to extract following the extraction of its overlying blocks represented by a set $\mathbb{O}_b$. $\mathbb{S}$ denotes a set of scenarios that quantify the joint uncertainty in grade, geometallurgical properties and material types. $\mathbb{T}$ represents the set of scheduling periods. Extracted material from mines can either be stockpiled, processed after crushing, or sent to waste. $v_{a,i,t,s}$ represents the amount of property a at location i in period t and scenario s. $\mathcal{S}$ represents the set of stockpiles. $SC_{i,a,t}$ represents the stockpiling cost of stockpile i for property a in period t. $RH_{i,a,t}$ represents the cost of re-handling material from stockpile i for property a in period t. $\mathcal{S}^c$ represents the set of crushers. $CR_{i,a,t}$ represents the crushing cost of material with crusher i for property a in period t. $\mathcal{P}$ represents the set of processing mills. $P_{i,a,t}$ represents profit (selling price − selling cost) for recovered product a with processing destination i in period t. $PC_{i,a,t}$ represents the processing cost with processing destination i for property a in period t. Properties such as metal tonnage, rock tonnage, and ore tonnage are represented by $\mathbb{P}$, and are calculated by adding the amount of metal, rock, and ore tonnage processed at different locations in a mining complex. Properties such as head grade copper, head grade arsenic, etc., are calculated thus: head grade copper is equal to total copper metal tonnage / total ore tonnage, head grade arsenic is equal to total arsenic metal tonnage / total ore tonnage at different locations. $\mathbb{H}_p$ represents the amount of different products recovered.

Production targets are usually present in a mining complex, such as (i) capacity targets represented by a set $\mathbb{P}_c$, (ii) quality targets represented by a set $\mathbb{H}_q$, and (iii) geometallurgical targets represented by a set $\mathbb{H}_g$. Mineability targets are also included in the model. Mineability targets ensure that the production schedules are practically mineable. N_b is the specified mining width. $w_{b,t}$ is the number of blocks scheduled in multiple periods inside a mining width. c_t^{smooth} is the penalty cost associated with not scheduling blocks inside the mining width in the same mining period. $c_{a,i,t}^+$ and $c_{a,i,t}^-$ represent the penalty costs for deviations from maximum/upper and minimum/lower production targets respectively for property a in period t. Profits and costs are discounted using an economic discount factor d, as in $P_{i,a,t} = P_{i,a,1}/(1+d)^t$. Costs of deviation from capacity, blending, geometallurgical and mine ability targets are also discounted using a geological discount rate r, as in $c_{a,i,t}^+ = c_{a,i,1}^+/(1+r)^t$ used from Dimitrakopoulos and Ramazan (2005), Ramazan and Dimitrakopoulos (2013), and Spleit and Dimitrakopoulos (2017). Geological risk discounting helps to defer the risk of not meeting production targets to later years when more information will be available.

2.2.2 Decision variables

Decision variables for the model are: (i) mine extraction sequence variables, $x_{b,t} \in \{0,1\}$, which define whether (1) or not (0) a block $b \in \mathbb{B}_m$ is extracted in period $t \in \mathbb{T}$, (ii) processing stream utilization variables, $y_{i,j,t,s} \in [0,1]$, which define the proportion of material sent from location i to subsequent location j in period t and under scenario s, and (iii) cluster destination policy variables, $z_{c,j,t} \in$

{0,1}, which define whether (1) or not (0) a cluster $c \in \mathcal{C}$ is sent to one of the allowed destinations j in period t. Cluster destination policies are based on clusters defined over multiple elements of interest (Goodfellow and Dimitrakopoulos, 2016). Continuous variables that allow deviations from different targets are also included in the model. Surplus variables, $d^+_{a,i,t,s}$, represent excesses from maximum target $U_{a,i,t}$ for property a in period t and scenario s. Shortage variables, $d^-_{a,i,t,s}$, represent shortages from minimum target $L_{a,i,t}$ for property a in period t and scenario s.

2.2.3 Objective function

The objective function (Eq. 2.1) of the model is a two-stage function that maximizes the value of the product generated from a mining complex and delivered to customers or spot markets, while minimizing the deviations from capacity, blending, geometallurgical and mineability targets, under grade, geometallurgical and material type uncertainties.

$$
\max \frac{1}{|\mathbb{S}|} \left\{ \sum_{s\in\mathbb{S}} \sum_{t\in\mathbb{T}} \left(\underbrace{\sum_{i\in\mathcal{P}} \sum_{a\in\mathbb{H}_p} P_{i,a,t} \cdot v_{a,i,t,s}}_{\text{Part I}} - \underbrace{\sum_{i\in\mathcal{P}} \sum_{a\in\mathbb{P}} PC_{i,a,t} \cdot v_{a,i,t,s}}_{\text{Part II}} \right. \right.
$$

$$
- \underbrace{\sum_{i\in\mathcal{S}^c} \sum_{a\in\mathbb{P}} CR_{i,a,t} \cdot v_{a,i,t,s}}_{\text{Part III}} - \underbrace{\sum_{i\in\mathcal{S}} \sum_{a\in\mathbb{P}} SC_{i,a,t} \cdot v_{a,i,t,s}}_{\text{Part IV}} - \underbrace{\sum_{i\in\mathcal{S}} \sum_{a\in\mathbb{R}} RH_{i,a,t} \cdot v_{a,i,t,s}}_{\text{Part V}}
$$

$$
- \underbrace{\sum_{i\in\mathcal{P}} \sum_{a\in\mathbb{P}_c} \left(c^+_{a,i,t} \cdot d^+_{a,i,t,s} + c^-_{a,i,t} \cdot d^-_{a,i,t,s} \right)}_{\text{Part VI}} - \underbrace{\sum_{i\in\mathcal{P}} \sum_{a\in\mathbb{H}_q} \left(c^+_{a,i,t} \cdot d^+_{a,i,t,s} + c^-_{a,i,t} \cdot d^-_{a,i,t,s} \right)}_{\text{Part VII}}
$$

$$-\sum_{i \in \mathcal{P}} \sum_{a \in \mathbb{H}_g} \underbrace{\left(c^+_{a,i,t} \cdot d^+_{a,i,t,s} + c^-_{a,i,t} \cdot d^-_{a,i,t,s} \right)}_{\text{Part VIII}} - \sum_{i \in \mathcal{S}^c} \sum_{a \in \mathbb{P}_c} \underbrace{\left(c^+_{a,i,t} \cdot d^+_{a,i,t,s} + c^-_{a,i,t} \cdot d^-_{a,i,t,s} \right)}_{\text{Part IX}}$$

$$-\sum_{i \in \mathcal{S}} \sum_{a \in \mathbb{P}_c} \underbrace{\left(c^+_{a,i,t} \cdot d^+_{a,i,t,s} + c^-_{a,i,t} \cdot d^-_{a,i,t,s} \right)}_{\text{Part X}} - \sum_{i \in \mathbb{M}} \sum_{a \in \mathbb{P}_c} \left(\underbrace{c^+_{a,i,t} \cdot d^+_{a,i,t,s}}_{\text{Part XI}} + \underbrace{c^-_{a,i,t} \cdot d^-_{a,i,t,s}}_{\text{Part XII}} \right) \Big\}\Big\}$$

$$-\sum_{t \in \mathbb{T}} \sum_{m \in \mathbb{M}} \sum_{b \in \mathbb{B}_m} \underbrace{\left\{ MC_{b,t} \cdot x_{b,t} + c^{smooth}_t \cdot w_{b,t} \right\}}_{\text{Part XIII}} \qquad (2.1)$$

Part I in the objective function represents the profits of different products produced and sold to the customers. Part II is the processing cost of the material at the different processing destinations. Part III represents the crushing cost at the different crushers. Part IV represents the stockpiling cost at the different stockpiles. Part V represents the cost of re-handling material from the different stockpiles. Parts VI, VII, and VIII represent the costs of deviation from the capacity, quality and geometallurgical targets at the processing streams. Part IX is the cost of deviation from the crusher capacity target. Part X is the cost of deviation from the stockpile capacity target. Part XI represents the cost of deviation from the mining capacity target. Part XII represents the mining cost at the different mines. Part XIII is the cost associated with the smoothness of schedules.

2.2.4 Constraints

The processing mill in the mining complex has multiple operating modes. Operating modes are defined herein as the configuration of the processing mill that

73

determines the grinding size of the materials. For example, fine grinding compared to coarse grinding operating mode. Different operating modes have different costs, energy consumption, and throughput. However, the cost, energy consumption, and throughput for a fixed operating mode might change based on the hardness of the material processed. For example, the fine grinding of hard material will require a longer residence time when compared to softer material, which results in higher costs and lower throughput. Uncertainty about the hardness of material results in inconsistent throughput, energy consumption, and operating costs with a fixed operating mode at the processing mill. Hardness indexes such as the SPI and BWI are non-additive properties which cannot be directly included in the outlined optimization model. A different approach is utilized in this work to overcome such non-additive issues with hardness indexes. First, the operating mode is fixed and linked to the proportion of hard and soft material that the processing mills can process to achieve maximum throughput and recovery, and minimum cost. Such proportions of hard and soft material are defined as geometallurgical targets $\mathbb{H}_g$ with the different processing mills for a fixed operating mode. Further, the material at the mine is characterized as soft and hard material based on its hardness indexes. For instance, if the hardness indexes of the mining block are above a certain threshold, it is defined as hard.

A certain amount of hard and soft material is extracted from the mine and sent to different processing streams, and is finally processed at the different processing mills. At the different processing mills, the total amount of hard and soft material is determined, and their ratio is calculated as $\left(\frac{amount\ of\ hard\ material}{amount\ of\ soft\ material} \right)$. Finally, deviations from the upper and lower limits of predefined geometallurgical targets $\mathbb{H}_g$,

74

are calculated using Equations 2.2 and 2.3. Such deviations are penalized in the objective function with a penalty cost. Penalizing such deviations ensures that the proportion of hard and soft material processed at the processing mills is within the required geometallurgical targets.

$$v_{h,i,t,s} - d^+_{h,i,t,s} \leq U_{h,i,t} \quad \forall\, h \in \mathbb{H}_g, i \in \mathcal{P}, t \in \mathbb{T}, s \in \mathbb{S} \tag{2.2}$$

$$v_{h,i,t,s} + d^-_{h,i,t,s} \geq L_{h,i,t} \quad \forall\, h \in \mathbb{H}_g, i \in \mathcal{P}, t \in \mathbb{T}, s \in \mathbb{S} \tag{2.3}$$

$$N_b \cdot x_{b,t} - \sum_{n \in N_b} \sum_{t' \in t} x_{n,t'} \leq w_{b,t} \tag{2.4}$$

Equation 2.4 calculates the number of blocks scheduled in multiple periods, $w_{b,t}$, inside a mining width of N_b and penalizes it with a cost of c_t^{smooth} in the objective function. Other constraints implemented in the model are capacity/quantity, blending/quality, destination policy, processing stream flow, reserve, and slope constraints as detailed in Goodfellow and Dimitrakopoulos (2016).

2.2.5 Solution approach

The simultaneous stochastic optimization model of mining complexes outlined in the previous sections is a large combinatorial optimization model with millions of binary decision variables. Such large optimization models cannot be solved with general purpose commercial solvers such as CPLEX. Metaheuristic algorithms have proven effective in generating good solutions in a reasonable computational time. The solution approach used to solve the model is the same as mentioned in Goodfellow and Dimitrakopoulos (2016, 2017). A brief review of the solution approach is outlined in Appendix 2.2.

2.3   Application at a large copper-gold mining complex

This section outlines the application of the method presented in the previous section at a large copper-gold mining complex and demonstrates its applied aspects. The management of the mining complex makes decisions based on a given strategic mine plan. The strategic mine plan currently used at the mining complex is developed with high-end conventional technologies as outlined in Section 2.3.3. The framework proposed herein generates a stochastic strategic mine plan that accounts for variability and uncertainty in the supply of materials and jointly optimizes the different components of the mining complex. The following sections compare and highlight the differences between the forecasted performance of the conventional and stochastic mine plans. The key aspect of the comparison is to highlight that the decision makers will make different decisions with the conventional mine plan as compared to the stochastic mine plan.

2.3.1   Overview of the copper-gold mining complex

The copper-gold mining complex consists of two mines (Mine A and Mine B) with 120,000 and 78,000 blocks, respectively, measuring 25x25x15 m^3 in size. The two mineral deposits consist of 8 different mine zones, 8 different alterations and 10 lithologies each (as provided by the mining company operating the mining complex). The stratigraphic sequence of material is waste rock, followed by copper oxide, mixed and copper sulphide. The mining complex produces copper concentrate and copper cathode as primary products, and gold, silver and molybdenum concentrates as secondary products. The material extracted from both mines is classified into 12 different material types for each mine and can be sent to one out of 9 different

destinations (5 crushers, 2 stockpiles, 1 bio leach pad (sulphide leach pad) and 1 waste dump) as shown in Figure 2-1. Material from five different crushers is further sent to three different processing mills that supply material to the port, and an oxide leach pad that supplies material to a copper cathode plant. The primary product generated at the port is copper concentrate. In addition, different secondary products such as gold concentrate, silver concentrate, and molybdenum concentrate are also generated at the port.

Figure 2-1 Flow of material at the copper-gold mining complex

The copper cathode plant generates copper cathode as the product. The port and the copper cathode plant produce the final products of the mining complex which are transported and sold to different customers. Geometallurgical targets related to the SPI and BWI hardness indexes are present with the three processing mills in the

mining complex. Different blending and capacity targets are also present with the different components of the mining complex.

2.3.2   Stochastic optimization

The stochastic optimization model generates the stochastic mine production plan which is referred to herein as the "stochastic plan". The different parameters associated with the stochastic plan are discussed in Sections 2.3.2.1 to 2.3.2.4.

2.3.2.1 *Stochastic orebody simulations*

Fifteen stochastic simulations are generated for each mine that quantify uncertainties in grade, geometallurgical properties and material types, using the direct block simulation with minimum/maximum autocorrelation factors (Boucher and Dimitrakopoulos, 2009). A brief review of the above simulation method can be found in Appendix 2.1. Grade uncertainty is quantified for seven different correlated properties, namely copper soluble (CuS), copper total (CuT), iron (Fe), arsenic (As), gold (Au), silver (Ag), and molybdenum (Mo). The major elements CuS, CuT, Fe, and As with higher sampling density were decorrelated and simulated together. The minor elements Au, Ag, and Mo with lower sampling density were decorrelated and simulated together.

The two mines have 8 different mine zones each, thus the simulation of grade properties were generated separately for each mine zone of each mine, each time using the samples of the seven correlated grade properties within the corresponding mine zone. The direct block simulation method was used to generate the correlated grade properties of the mining blocks within each mine zone. The combination of the simulated mining blocks from different mine zones provided the final simulation of

78

correlated grade properties for each mine. Figure 2-2 represents the three randomly selected simulations for the two deposits for the copper total grade property compared to the estimated deposit. Variability in the copper total grade property can be easily perceived from the simulations compared to a smooth representation of the estimated deposits. Geometallurgical uncertainty is quantified for two hardness index values, the SPI and BWI. The two mines have 10 different lithologies each, thus the simulations of geometallurgical properties were generated separately for each lithology of each mine, each time using the samples of the SPI and BWI within the corresponding lithology. The direct block simulation method was used to generate the correlated geometallurgical properties of the mining blocks within each lithology. The combination of the simulated mining blocks from the different lithologies provided the final simulation of the correlated geometallurgical properties for each mine.

Figure 2-2 Simulation of copper total for the two deposits compared to the estimated deposits

Figure 2-3 represents the three randomly selected simulations for the two deposits for the SPI hardness index compared to the estimated deposits. A large variability in hardness properties is perceived from the simulations compared to a very smooth representation from the estimated deposits. Variability and uncertainty in the grade and geometallurgical properties result in an uncertainty in material types which is shown in Figure 2-4. Materials are first classified as sulphide, oxide, waste and mixed based on geological properties (alteration and lithology) and grade properties (copper soluble, copper total, and copper soluble to copper total ratio), then geometallurgical properties (SPI and BWI) are used to further characterize the materials as soft or hard. Such classification strategies result in twelve different material types with distinct geological, grade, and geometallurgical properties. Material type uncertainty is quantified for the twelve different material types (i) soft waste, (ii) hard waste, (iii) mixed, (iv) high grade oxide, (v) low grade oxide, (vi) soft high grade sulphide, (vii) hard high grade sulphide, (viii) soft low grade sulphide, (ix) hard low grade sulphide, (x) soft sulphide ore, (xi) hard sulphide ore, and (xii) oxide ore. A comparision between the histograms of the estimated and simulated SPI models is presented in Appendix 2.3

Figure 2-3 Simulation of SPI for the two deposits compared to the estimated deposits

Figure 2-4 Uncertainty in material types for the two deposits compared to the estimated deposits

2.3.2.2 *Economic and operational parameters*

The economic and operational parameters used for the copper-gold mining complex are outlined in Table 2-1 and Table 2-2, respectively. The parameters are scaled values for reasons of confidentiality.

Table 2-1 Parameters used for the mining complex

Attribute	Value
Economic discount rate	8%
Geological discount rate	10%
Gold, silver, and molybdenum price (US M$/Metal Ton Recovered)	36, 0.44 and 0.011
Copper price (US $/Metal Ton Recovered)	4960
Mining cost {excluding hauling cost} (US $/Rock Ton)	0.60
Hauling cost based on depth (US $/Rock Ton)	Mine A: 0.4 to 1.27 Mine B: 0.52 to 1.09
Milling cost including crushing cost (US $/Ore Ton)	5.7
Cost for re-handling material from stockpile (US $/Ore Ton)	0.18
Oxide leach pad leaching cost including crushing cost (US $/ Ore Ton)	5.7
Bio leach pad leaching cost (US $/Ore Ton)	1.65
Selling Cost (US $/Metal Ton Recovered)	Mill -514 Oxide leach pad - 496 Bio leach pad - 496

Table 2-2 Operational parameters used for the mining complex

Attribute	Value
Blocks for Mine A	120634
Blocks for Mine B	78355
Scheduling period	8 years
Number of clusters of the different material types	25
Mining width	200 m
Slope angle for Mine A	37
Slope angle for Mine B	45

Recovery of copper at mills (Mill1, Mill2, and, Mill3)	0.80, 0.83 and 0.82
Recovery for minor elements at processing mills	0.20
Recovery of copper at oxide leach pad	0.65
Recovery of copper at bio-leach pad	0.27

2.3.2.3 *Production targets*

Figure 2-5 and Figure 2-6 represent the different yearly quality, quantity, and geometallurgical targets of the different processing streams in the mining complex. The maximum capacity in these figures represent the upper limit capacity constraint. The color legend represents the different rocktypes in the two mines, and the different processing streams have an associated color band that shows what rocktypes are eligible to be processed at that processing stream.

Figure 2-5 Yearly quantity and quality targets with crushers, stockpiles, and leach pads used for the mining complex

Mine A and Mine B	Mill 1, Mill 2, and Mill 3
Waste (Hard, Soft)	• Max Capacity Mill 1 – 28.1 %
Mixed	• Max Capacity Mill 2 – 33.6 %
Low Grade Oxide	• Max Capacity Mill 3 – 38.3 %
High Grade Oxide	• Max Capacity Total Mill – 100 %
Low Grade Sulphide (Hard, Soft)	• Max Fe head grade (%)– 100 %
	• Max As head grade (PPM) – 100 %
High Grade Sulphide (Hard, Soft)	• Max CuT/Fe ratio – 100 %
Sulphide Ore (Hard, Soft)	• Hard/Soft ratio Mill 1 – [0.5,1]
	• Hard/Soft ratio Mill 2 – [1,1.6]
Oxide Ore	• Hard/Soft ratio Mill 3 – [1.2-1.8]

Figure 2-6 Yearly quantity, quality, and geometallurgical targets with processing mills used for the mining complex

The geometallurgical targets mentioned in Figure 2-6 were selected based on the information provided by the mining company operating the mining complex. The mining complex had selected such targets based on studies about the configuration and operating modes of their different processing mills to achieve optimal throughputs. The modelling of hardness properties can also be integrated in the model by associating materials to throughput and cost based on their hardness properties. However, the approach outlined in this work is selected to adhere to the specification and targets provided for the specific case study and the given mining complex. The different quantity and quality targets are scaled values for reasons of confidentiality.

Penalty costs associated with the violation of different production targets with the different components of the mining complex are summarized in Table 2-3. The penalty costs mentioned in Table 2-3 were applied on the absolute amount of deviation from the different production targets. However, the key aspect to note is that a

deviation of 1 ton from the processing mill tonnage target will not have the same effect as a deviation of 1 ppm from the processing mill arsenic grade target. Therefore, the penalty costs are selected to reflect the magnitude of the penalty value, which would be incurred with 1 unit of deviation from the related production target. For example, as shown in Table 3, the processing mills have a penalty cost of 100 US $/Ton for 1 unit of deviation from the tonnage target, compared to a penalty cost of 0.1 billion US $/PPM for 1 unit of deviation from the arsenic grade target. Similarly to other input parameters, the penalty costs are decided based on trial and error with consideration of the magnitude of different production targets to manage and minimize the technical risk for the different production targets. The trial and error approach entailed changing the different penalty costs while accounting for the magnitude aspect described above and observing the amount of deviations from the different production targets in the mining complex, finally selecting the penalty costs which provided the minimum technical risk for the different production targets. Additional discussion on penalty costs may be found in Benndorf and Dimitrakopoulos (2013).

Table 2-3 Penalty costs used for the mining complex

Targets	Penalty Cost
Capacity (C1, C2, C3, C4, C5)	40, 40, 40, 40, and, 40 US $/Ton
Capacity (Mill1, Mill2, Mill3)	100,100, and,100 US/Ton
Mining width constraint	50000 and 50000 US $
Stockpile Capacity (Oxide, Sulphide)	40 and 40 US $/Ton
Leach pad Capacity (Oxide, Sulphide)	40 and 40 US $/Ton
Ratio (CuT/Fe, CuS/CuT, hard/soft)	1 billion US $/unit
Fe Grade, and As Grade	0.1 billion US $/%, and 0.1 billion US $/PPM

2.3.2.4 *Metaheuristic parameters*

The solution approach outlined in Section 2.2.5 for solving the simultaneous stochastic optimization model of mining complexes requires different parameters which are summarized in Table 2-4. These parameters are selected by trial and error using the following criterion: they must generate a stable solution that maximizes the NPV while meeting production targets. The model has 1.61 million binary and 50,000 continuous decision variables. The solution time for the model is around 72 hours.

Table 2-4 Metaheuristic parameters used in the solution approach for the mining complex

Parameter	Value
Initial annealing temperature	0.5
Cooling factor	0.95
Cooling after perturbations	2000
Perturbations before diversification	120,000
Number of diversifications	13

2.3.3 Conventional optimization

The long-term mine production plan currently used at the mining complex is optimized using a two-step optimization approach: (a) the extraction sequences of multiple mines are optimized independently of each other using Whittle version 4.5.4 (Gemcom, 2005; Dassault, 2017), a widely used software for strategic mine planning, (b) the destination of the extracted material is decided based on cut-off grade policy presently used at the mining complex and is based on Lane's approach (Lane, 1988; Rendu, 2014; Khan and Asad, 2020). Then, the utilization of different processing streams is defined using a separate optimization model. In addition, this two-step optimization process is performed over the estimated mineral deposits, as is the

standard practice in the mining industry. This long-term mine production plan of the copper-gold mining complex generated with this two-step approach is referred to herein as the "conventional plan". Economic parameters, operational parameters and production targets used to generate the conventional plan are the same as the ones used for the stochastic plan. The different geometallurgical and quality targets are only used during the optimization of the processing stream utilization decisions in the conventional mine plan. The extraction sequence and cut-off grade policy in the conventional mine plan accounts only for the quantity targets. Estimated mineral deposits used in the conventional plan are shown in Figure 2-2 and Figure 2-3 for grade and geometallurgical properties respectively. Material types of the estimated mineral deposits are shown in Figure 2-4.

2.3.4 Results of the stochastic optimization and comparison to conventional optimization

The results of the stochastic optimization of the copper-gold mining complex are presented in this section. Results are reported using the 10th, 50th, and 90th percentile risk profiles (P10, P50, and P90, respectively) of the different performance indicators with respect to the simulated orebody models of the two mines. P10 stands for a 10 % probability that the values are below this value, P50 is 50 % probability of having values below this value and, similarly, P90 stands for 90 % probability of having values below this value. The forecasts of the stochastic plan are compared to the forecast of the conventional plan throughout its presentation and discussion, to highlight the differences in forecasted performance between the two mine plans and to show the added value of the stochastic framework, where appropriate.

87

2.3.4.1 *Capacity targets*

Figure 2-7(a)-(h) presents the risk profiles of meeting the capacity target for crushers 1, 2, 3, 4 and 5, and mills 1, 2, and, 3. Figure 2-7(a) shows that the stochastic plan better utilizes and respects the crusher capacity (P10, P50, and P90 of the risk profiles are very close to the capacity target). The capacity target is only violated slightly in year 7, but is, overall, well controlled throughout the scheduling period in the stochastic plan (Figure 2-7(a)). Similar behavior is observed for the stochastic plan for crusher 2 (Figure 2-7(b)), crusher 3 (Figure 2-7(c)), and crusher 5 (Figure 2-7(e)), where the capacity target is well respected and utilized over all the scenarios and throughout the scheduling period. However, the conventional plan shows very low production in year 6 for crusher 1, crusher 2, and crusher 5. In addition, the production for crusher 2 and crusher 3 is lower throughout the scheduling period in the conventional plan (Figure 2-7(b) and (c) respectively) compared to the stochastic plan.

Figure 2-7 Forecasts for crushers 1 to 5 and capacity targets for the stochastic and conventional plans for mills 1 to 3

Crusher 4 (Figure 2-7(d)) shows only a small violation of the capacity target in the stochastic plan in year 2, but it utilizes the capacity of crusher 4 better than the conventional plan. However, the risk profile for the stochastic plan for crusher 4 has

a much wider envelope compared to the other crushers. Similar behavior is present for mill 1 (Figure 2-7(f)) and mill 2 (Figure 2-7(g)) for the stochastic plan. Such a wide risk envelope in the capacity target of crusher 4, mill 1 and mill 2 originates from the processing of materials from both the mines, which results in relatively high grade, geometallurgical and material type uncertainties. The high uncertainties originate from the combination of supply uncertainty scenarios from multiple mines. In the conventional plan, the capacity target is respected for crusher 4 (Figure 2-7(d)) but does not utilize the maximum crusher capacity. Similarly, the mill capacity is not fully utilized in the conventional plan for mill 1 and mill 2 with almost negligible production in year 6 for mill 1 (Figure 2-7(f)) and mill 2 (Figure 2-7(g)). The capacity target of mill 3 (Figure 2-7(h)) has a relatively narrow risk envelope compared to the other mills because it processes material from only one mine. The stochastic plan better utilizes the capacity target in mill 3 compared to the conventional plan, where the production is almost negligible for the conventional plan in years 1, 6, and 7 (Figure 2-7(h). Both the conventional and stochastic plans respect the capacity target of all the mills throughout the scheduling period.

Figure 2-8 Forecasts for sulphide stockpile, oxide stockpile, oxide leach pad, bio-leach pad capacity target, and mining rate of mines A and B for the stochastic and conventional plans

Figure 2-8(a)-(d) presents the risk profiles of the capacity target for the oxide leach pad, bio-leach pad (sulphide leach pad), sulphide stockpile and oxide stockpile. A small violation is only observed in year 2 in the stochastic plan for the oxide leach pad capacity target (Figure 2-8(a)) with a wide risk envelope originating from the reason explained above. The capacity target is well respected for the bio-leach pad (Figure 2-8(b)), stockpile sulphide (Figure 2-8(c)), and stockpile oxide (Figure 2-8(d)) in the stochastic and conventional mine plans. Figure 2-8(e)-(f) present both the

91

forecasted mining rates of the stochastic and conventional mine plans for mines A and B, respectively. The mining rates and, consequently, the mine production plans of the two mines are very different; more specifically, the stochastic mine plan extracts more materials from mine A and less materials from mine B during the initial periods, when compared to the conventional mine plan.

2.3.4.2 *Blending targets*

Figure 2-9(a)-(d) represents the risk profiles for the blending target of the copper total to iron ratio at the different processing mills, and the ratio of copper soluble to copper total at the bio-leach pad. The blending target is very well respected for the stochastic plan for mill 1 (Figure 2-9(a)), mill 2 (Figure 2-9(b)), and mill 3 (Figure 2-9(c)) with minor violations in years 2 and 4. The conventional plan shows larger violations from such targets for all the mills. The blending target of the bio-leach pad is very well respected in the stochastic and conventional plans. Figure 2-10(a), (c), and (e) show the risk profile of the iron (Fe) head grade target for the different processing mills. The stochastic plan has a small violation for the iron head grade target for mill 1 (Figure 2-10(a)) in year 6. Mill 2 (Figure 2-10(c)) presents a small violation in year 6 for the iron head grade target in the stochastic plan. Mill 3 (Figure 2-10(e)) presents no violation for the iron head grade target in the stochastic plan. The conventional plan also presents a violation from year 6 onwards for mill 1 (Figure 2-10(a)) for the iron head grade target. The conventional plan shows no violation for the iron head grade target for mill 2 (Figure 2-10(c)) and mill 3 (Figure 2-10(e)).

Figure 2-9 Forecasts of blending target (CuT/Fe and CuS/CuT ratio) for mill 1, 2, and 3 and bio-leach pad for the stochastic and conventional plan

Figure 2-10(b), (d), and (f) show the risk profile of the arsenic (As) head grade target for the different processing mills. The stochastic plan respects the arsenic head grade target very well for mill 1 (Figure 2-10(b)), mill 2 (Figure 2-10(d)), and mill 3 (Figure 2-10(f)). In the conventional plan, a large violation from the arsenic head grade target is expected over the scheduling period, particularly in years 2,3,6,7, and 8 for mill 1 (Figure 2-10(b)). Similar behavior is observed for the arsenic grade target of mill 2 (Figure 2-10(d)), with a small violation in the initial years and a larger one in the later years. The arsenic target is well respected in the stochastic and conventional plans for mill 3 (Figure 2-10 (f)).

Figure 2-10 Forecasts of blending target (Fe and As head grade) for mill 1, 2, and 3 for the stochastic and conventional plan

2.3.4.3 *Geometallurgical targets*

Figure 2-11(a), (b), and (c) display the risk profiles of geometallurgical targets (hard/soft material ratios) for mill 1, mill 2, and mill 3. The stochastic plan only presents a very small violation from geometallurgical targets for mill 1 as compared to the conventional plan, which presents a very large violation from such targets for mill 1 and mill 2 (Figure 2-11(a) and (b) respectively). The conventional plan does not account for geometallurgical properties in the optimization model but only uses it

for determining processing stream utilization decisions. The geometallurgical targets are well respected for mill 3 for the stochastic and conventional plans (Figure 2-11(c)). In addition, the stochastic plan shows stable geometallurgical target forecasts with significantly less fluctuation over the scheduling period as compared to a higher fluctuation in the conventional plan. Maintaining such a stable proportion of hard and soft material helps to achieve better throughput rates and efficient planning of processing mill operations.

Figure 2-11. Forecasts of geometallurgical target for mill 1, 2, and 3 for the stochastic and conventional plan

2.3.4.4 *Metal production*

Figure 2-12(a)-(d) shows the risk profiles of the recovered primary (copper concentrate) and secondary products (gold concentrate, silver concentrate, and molybdenum concentrate). The values are scaled with respect to the conventional plan (the conventional plan is 100 %) for reasons of confidentiality. The stochastic plan

produces 12.5 % higher primary copper product (Fig. 12(a)) as compared to the conventional plan over the scheduling period of 8 years.

Figure 2-12 Forecasts of cumulative copper, gold, silver, and molybdenum recovered in the mining complex for the stochastic and conventional plan

A higher production of secondary products, that is (i) 22.9 % of gold (Figure 2-12(b)), (ii) 32.4 % of silver (Figure 2-12(c)), and (iii) 34.7 % of molybdenum (Figure 2-12(d)), is observed in the stochastic plan compared to the conventional plan. Such significant differences come from (i) capitalizing on the synergies of simultaneous optimization of a complete mining complex, (ii) capitalizing on the variability of material supply, and (iii) capitalizing on the added value of secondary products to drive the optimization process. Note that the conventional plan only considers copper in its production planning because the mining complex is a major producer of copper products and, therefore, does not account for other minor elements

in planning decisions. As a result, the conventional plan is unable to capitalize on benefits generated from the added value of such secondary products to drive the optimization process. The secondary products are recovered as by-products and, therefore, the value from such products are included in the forecast of the conventional plan during the comparison presented herein.

2.3.4.5 *Discounted cash flows*

Figure 2-13 represents the risk profile for NPV forecasts. NPV is the cumulative discounted cash flow. The values are scaled with respect to the forecast of the conventional plan (the conventional plan is 100 %) for reasons of confidentiality. The stochastic plan presents a 19.3 % higher NPV compared to the conventional plan, for several compounding reasons. The large difference in the NPV forecast of the stochastic plan compared to the forecasts of the conventional plan is due to:

(i) The ability to capitalize on synergies between the different components of the mining complex by jointly developing the extraction sequence, destination policies, and processing stream utilization under grade, geometallurgical and material type uncertainties, to maximize the value of products sold to the different customers.

(ii) Incorporating uncertainty and variability of grade, material type and geometallurgical properties of the material in the stochastic optimization model , which helps to manage the technical risk throughout the life of the mining complex and increases the expectation of meeting the different production targets, including major effects on blending, destination policies, and processing stream utilization.

(iii) The integration of different blending and capacity restrictions in the simultaneous stochastic optimization model, which allows for better

97

utilization of processing mills and crushers through the blending of material from multiple mines (as observed in Figure 2-7). The conventional plan, on the other hand, optimizes the extraction sequence of multiple mines independently and, then, uses a separate optimization model to define processing stream utilization decisions to satisfy blending restrictions. Two-step optimization leads to the underutilization of processing mills and crushers to meet the blending restriction.

(iv)   The ability to focus on the value of product sold rather than the economic value of the mining blocks

(v)   The utilization of secondary products such as gold, silver, and molybdenum in the optimization process of the long-term stochastic mine plan (as observed in Figure 2-12).

Figure 2-13 Cumulative discounted cash flow forecasts for the stochastic and conventional plan

2.3.5   Production schedule

Figure 2-14 displays the cross-sectional and top views of the extraction sequence of the copper-gold mining complex. The smoothing constraints implemented with the

model help to generate mineable shapes for the extraction sequence in the stochastic plan. The stochastic plan extracts material in different proportions form the multiple mines compared to the conventional plan, to meet the required production targets. In addition, the stochastic and conventional schedules are physically different, with different areas being mined in different years in the two schedules.

Figure 2-14 Generated schedule (a) top view mine A, (b) cross-section view mine A, (c) top view mine B, (d) cross-section view mine B; for the stochastic and conventional plan

2.4   Conclusions

The simultaneous stochastic optimization of mining complexes is revisited and extended to include geometallurgical uncertainty and decisions. The model presented herein is very flexible and can be applied at any mining complex with different geometallurgical properties, production characteristics and constraints. The parameters of the stochastic optimization model presented require some trial and error in generating mine plans that maximize NPV and generates stable solutions, similarly

to any other optimization method for strategic mine planning. The contribution and applied aspects of the proposed approach are highlighted with an application at a large copper-gold mining complex that consists of 2 mines, 5 crushers, 2 stockpiles, 3 processing mills, 1 waste dump and 2 leach pads, and produces copper, gold, silver and molybdenum products for different customers and spot markets. Two geometallurgical properties related to the grindability of the material, the SPI and BWI, are considered in the optimization model, which increases the utilization of different processing mills in the mining complex. The forecasts of the stochastic mine production plan show an increased expectation of meeting the different production targets, including consistency in throughput and hardness of material processed at the processing mills and crushers, when compared to the forecasts of the conventional mine production plan. In addition, the forecasts of the stochastic mine production plan show a 12.5% higher production of copper metal compared to the forecasts of the conventional mine production plan.

The utilization of secondary products in the stochastic optimization model increases the forecasts of production of gold, silver, and molybdenum by 22.9%, 32.4%, and 34.7% compared to the forecasts of the conventional mine production plan. The stochastic mine production plan also improves the NPV of the mining complex by 19.3% compared to the conventional mine production plan. Future work can consider integrating the decisions about operating modes such as coarse, medium, and fine grinding of material in the optimization model. Incorporating such operating mode decisions in the optimization model will help to enhance the performance of comminution circuits by choosing different operating modes based on the hardness of the material. Future directions also aim at integrating more geometallurgical properties in the optimization model directly instead of characterizing materials as hard and soft. The results from the case study indicate that the available facilities

100

(crushers, processing mills, leach pads, etc.) may be better utilized, providing further opportunities for value improvement through the inclusion of capital expenditure decisions in the simultaneous stochastic optimization model.

Appendix 2.1

This section briefly describes the direct block simulation method and the results of the simulation algorithm utilized to generate stochastic orebody models of the mineral deposits in the copper-gold mining complex. Interested readers are referred to Boucher and Dimitrakopoulos (2009) for more details.

Figure 2-15 Direction block simulation with minimum/maximum autocorrelation factor method

The direct block simulation method with a minimum/maximum autocorrelation factor requires the discretization of blocks into point support scale, transforming the correlated attributes from data space to normal score space, utilizing MAF to orthogonalize the correlated attributes into independent MAF factors, simulating MAF factors in point support scale, averaging point support scale values to generate block support scale values for further conditioning, concurrently back transforming

the point support simulated values to data space and averaging to generated simulated values at block support scale (Figure 2-15).

Appendix 2.2

This section briefly describes the metaheuristic solution approach used in the case study to solve the stochastic optimization model of mining complexes. The metaheuristic algorithm is the same used in Goodfellow and Dimitrakopoulos (2016):

$$P(g(\Phi), g(\Phi'), Temp) = \begin{cases} 1, & if \ g(\Phi) \leq g(\Phi') \\ \exp\left(-\dfrac{|g(\Phi') - g(\Phi)|}{Temp}\right), & otherwise \end{cases} \quad (2.5)$$

Metaheuristic approaches require an initial solution and therefore an initial solution is first generated using a greedy algorithm. The greedy algorithm generates a solution with the highest objective function value that satisfies the slope restrictions. The initial solution is then modified by a set of perturbations (neighborhood solution) to generate a new solution. Let Φ be the current solution with an objective value of $g(\Phi)$. The algorithm first groups the neighborhoods by the decisions they modify and then randomly selects one out of three decisions to modify: (i) the extraction sequence, $x \in \Phi$, (ii) the destination policy, $z \in \Phi$, and (iii) the processing stream utilization decision, $y \in \Phi$. Suppose that Φ' is the new neighborhood solution and is accepted or rejected using Equation 2.5. In the normal simulated annealing algorithm (Equation 2.5), a single temperature is used for a single neighborhood. In multi-neighborhood simulated annealing, a parameter $\delta \in [0,1]$ is defined and used, which

represents the probability of accepting a solution. The temperature $Temp$ for each decision is calculated using the parameter δ.

Each decision has a set of perturbations that can be used to modify the initial solution. For instance, the extraction sequence has perturbations such as (i) swapping the period of a block, (ii) changing the period of a block, (iii) changing the period of a block and all of its predecessors, etc.; processing streams have perturbations such as (i) changing the proportion of material at one or multiple processing streams, etc.; destination policies have perturbations such as (i) changing the destination of one or multiple clusters, etc. As the algorithm progresses, the probability of selecting the perturbations are modified and adapted, which is why it is known as an adaptive neighborhood search method.

Appendix 2.3

This section compares the simulated and estimated orebody models of the mineral deposits in terms of their grade-tonnage curves and histogram plots for estimated vs simulated values for the hardness index (SPI). Figure 2-16 represents the grade-tonnage curves for CuT for mines A and B. Figure 2-16 shows that the proportions of high- and low-grade materials are different in the estimated and simulated deposits. The smoothing of high values in the estimated deposits, as observed in Figure 2-2, is also shown in the grade-tonnage curves. Figure 2-17 shows the histograms of the SPI values of the estimated and simulated deposit models for mines A and B. Both Figure 2-17(a) and (b) show the effects of smoothness of the conventionally estimated deposits. The histograms are consistent with the conclusions drawn from Figure 2-3,

which also show the smoothed representation of SPI values in the estimated deposits compared to the variability of SPI values in the corresponding simulated models.

Figure 2-16 Grade-tonnage curve for copper total grade for (a) mine A and (b) mine B

Figure 2-17 Histograms of estimated and simulated hardness index (SPI) of the mineral deposits in (a) mine A and (b) mine B

CHAPTER 3
Adaptive Self-learning Mechanisms for Updating Short-term Production Decisions in an Industrial Mining Complex

The previous chapter proposed a new approach to integrate supply uncertainty related to geometallurgical properties in the simultaneous stochastic optimization of long-term production plan of a mining complex. This chapter develops a new self-learning framework to learn to respond to the incoming new information by adapting the short-term flow of materials in a mining complex within its long-term production plan.

3.1   Introduction

A mining complex is an integrated value chain network with multiple interlinked components including suppliers of raw materials (mineral deposits and external inventories), heavy earth moving equipment (shovel, trucks, and conveyor belts), handling facilities (crushers, stockpiles, and waste dumps), processing facilities (mineral processing mills and leach pads), and customer/commodity markets. Uncertainty is a characteristic of a mining complex, starting from the supply of different types of raw materials extracted from the mineral deposits involved (Dimitrakopoulos et al., 2002). Stochastic optimization models account for uncertainty and generate production decisions that yield higher value and manage the technical risk of not meeting the production targets (Matamoros and Dimitrakopoulos,

2016; Mai et al., 2019). A long-term production plan of a mining complex determines the annual strategic decisions that maximize net present value (NPV) and meets different production targets, while accounting for uncertainty in the supply of different types of materials (Montiel and Dimitrakopoulos, 2015, 2017, 2018; Goodfellow and Dimitrakopoulos, 2016, 2017). The short-term production plan determines the daily/weekly/monthly production decisions within the long-term production plan to meet annual targets. A review of short-term production planning in mining operations can be found in Blom et al. (2019). In addition to supply uncertainty, the short-term production plan accounts for uncertainty in the performance of equipment to determine the production decisions about the sequence of extracting materials from the mineral deposits, equipment assignment and allocation (Matamoros and Dimitrakopoulos, 2016; Quigley and Dimitrakopoulos, 2019), as well as the flow of materials from mineral deposits to customers and commodity markets. A major short-term production decision is to determine the flow of materials in a value chain that first includes deciding which handling facilities to send the extracted materials, often refered to as destination policies (Asad et al., 2016), and then involves determining how to utilize the processing facilities to produce the final products sold to customers/markets, often referred to as processing stream utilization.

New digital technologies, including the development of advanced sensors and monitoring devices, have enabled the acquisition of new information about the performance of the different components of a mining complex that affect the flow of materials in a value chain. Sensors installed on drills, shovels, trucks, conveyor belts, crushers, and mineral processing mills (Goetz et al., 2009; Dalm et al., 2014, 2018; Iyakwari et al., 2016; Wambeke and Benndorf, 2018) continuously measure the

performance of the mining equipment and processing streams (processing and handling facilities), as well as different pertinent properties of the materials being handled. In addition to the new sensor information, conventional sources of new information include blasthole sampling that determines the pertinent properties of materials extracted (Rossi and Deutsch, 2013), monitoring devices that measure the performance of equipment (Koellner et al., 2004), and tracking devices that track the location of materials (Brewer et al., 1999; Rosa et al., 2007).

Figure 3-1 The proposed continuous updating framework to adapt the short-term production plan of a mining complex with incoming new information

The core existing technologies can only integrate new information that is conventionally collected, such as grade control that integrates blasthole data to identify ore/waste boundaries in the blasted areas of mineral deposits (Verly, 2005; Dimitrakopoulos and Godoy, 2014) or dispatching stations that monitor the equipment for assignment and dispatch decisions (Kargupta et al., 2010; Nguyen and

Bui, 2015). However, these technologies are unable to integrate the sensor-generated information to adapt the short-term production plan. A continuous updating framework, shown in Figure 3-1, is needed to adapt the short-term production plan of a mining complex with new information generated from both sensors and conventional sources. The continuous updating framework consists of two parts. First, the new information generated from the different sources in a mining complex is used to update the performance of its different components, which includes uncertainty in the supply of materials from the mineral deposits, the performance of equipment, and the processing streams' capabilities (productivity, recovery, etc.). Second, the updated performance of the different components of a mining complex is then fed to an artificial intelligence framework, which, in the present work, is a neural network agent that is trained using policy gradient reinforcement learning to adapt the short-term production plan. The adapted short-term production plan is fed back to the mining complex to generate updated production forecasts. The adapted production plan is then followed, more sensor data is collected as the mining operations progress, and the production plan is adapted again, and the cycle continues. Benndorf and Buxton (2016) proposed a framework to update the mine planning decisions with new information. Related is also the work of Hou et al. (2015) and Shirangi (2017), who proposed a continuous updating framework to update the production plan of smart oil fields. However, the existing frameworks, both in mine planning and smart oil fields, require re-optimization of the production plan, which is computationally expensive with the available optimization techniques. Lamghari (2017) provided a detailed review of the different techniques used for production planning in mining complexes and smart oil fields. The new information generated in a mining complex can be

109

categorized as "soft" and "hard" data, based on the precision of their measurement. Sensor-generated information is "soft" data because it is noisy, uncertain, and ambiguous when collected during operations from different components of a mining complex. Direct measurements, such as those derived from drillhole samples, which are analyzed in geochemical laboratories and are substantially more precise, are considered "hard" data. Consequently, the first part of the continuous updating framework in a mining complex, as shown in Figure 3-1, aims at generating updated uncertainty models of the different components of a mining complex that are consistent with the hard data and minimize the mismatch between (a) the observed and forecasted production data, as well as (b) the soft and hard data. Evensen et al. (1994a) proposed the ensemble Kalman filter (EnKF) that updates the non-linear processes with new information and has long been used for petroleum reservoir flow simulation and production forecasting (Dovera and Della Rossa, 2011; Xue and Zhang, 2014; Kumar and Srinivasan, 2019; Xu and Hernández, 2019). The ensemble Kalman filter is a two-step assimilation process that first generates a model-based prediction based on initial simulations for a non-linear process and then corrects such predictions with new observed information. For example, in mining, observations made at the processing mills can be integrated to correct the initial simulated properties. The flow of materials from mine to mill is first tracked using sensors, then the initial simulated models are used to generate model-based prediction for the processing mill observations, and finally initial simulated properties are corrected based on the weighted difference between the model-based prediction and the observation. The method has been successfully applied to update pertinent attributes of mineral deposits (Benndorf, 2015; Dalm et al., 2018; Yüksel et al., 2018). Methods

such as randomized maximum likelihood (Sarma et al., 2006; Chen and Oliver, 2012; Vo and Durlofsky, 2014; Shirangi, 2017) and Markov mesh models (Panzeri et al., 2016) are also used to update the pertinent petroleum reservoir-related attributes. Vargas-Guzmán and Dimitrakopoulos (2002) and Jewbali and Dimitrakopoulos (2011) proposed a column-wise decomposition of the covariance matrix (CSSR) to update the pertinent attributes of mineral deposits with new hard data. However, the CSSR method cannot integrate the soft information generated from sensors. The outlined methods for updating pertinent attributes of mineral deposits with EnKF and CSSR are limited to a single attribute. This paper presents a new extension of EnKF that allows the updating of multiple correlated attributes in mineral deposits with minimum/maximum autocorrelation factors (Desbarats and Dimitrakopoulos, 2000).

The second part of the updating framework (Figure 3-1) aims at adapting the short-term production plan of a mining complex with the updated uncertainty models of its different components. Reinforcement learning methods are efficient in decision-making with new information. In recent years, reinforcement learning-based methods have shown exceptional performance at generating neural network agents that are capable of making very efficient decisions for different complex environments (Aissani et al., 2012; Mnih et al., 2013; Silver et al., 2016; Barde et al., 2019). Paduraru and Dimitrakopoulos (2018) proposed a Bayesian reinforcement learning algorithm to optimize the destination policies of materials in a mining complex. However, the method developed requires a predefined extraction sequence to calculate the expected a posteriori improvement in the objective function during the optimization. Paduraru and Dimitrakopoulos (2019) proposed a policy gradient reinforcement learning algorithm to optimize the neural network destination policies

of materials in a mining complex while accounting for supply and equipment performance uncertainties. The neural network destination policies increased the expected NPV by 6.5 % compared to the mine's cut-off grade destination policies for a copper mining complex. However, the method is (a) limited to a single product mining complex, and (b) does not provide a required continuous updating of the short-term production plan regarding destination policies of materials with the new information generated from sensors and/or conventional sources.

The work presented herein proposes a novel continuous updating framework that combines a new extension of the EnKF method and a policy gradient reinforcement learning method to adapt the short-term flow of materials in a multiple product mining complexes with incoming new information. The continuous updating framework allows a mining operation to learn, adapt, and make more informed short-term production planning decisions in real-time with incoming new information, allowing the operation to meet its production targets more closely. First, the proposed extension of the EnKF model is used to update the multiple pertinent correlated attributes in a mineral deposit with incoming new information. This part of the updating framework ensures that the ambiguous information is handled efficiently using Kalman gain in the proposed extension of the EnKF method. Second, the model presented in Paduraru and Dimitrakopoulos (2019) is further developed to account for multiple products in a mining complex. The second part of the updating framework uses an extraction and hauling simulator to generate samples for training the neural network destination policies agent through policy gradient reinforcement learning. In the following sections, the proposed continuous updating framework that adapts the short-term production plan in terms of the flow of materials with incoming new information is

112

detailed. Next, an application of the proposed continuous updating framework at a real copper mining complex is presented to show the efficiency and applied aspects of the proposed framework compared to the mine's cut-off grade destination policies. Conclusions and directions for future research follow.

3.2   Method

This section outlines the algorithm related to the two parts of the proposed continuous updating framework to update the short-term flow of materials in a mining complex with incoming new information. Please note that the notation used in the proposed framework is provided in Appendix 3.1.

3.2.1   Updating stochastic orebody simulations

The method proposed to update simulations of a mineral deposit with new information uses ensemble Kalman filter (EnKF) (Evensen, 1994a), which is modified to account for multiple correlated attributes. The group of simulations of mineral deposits is herein referred to as ensembles. The complete process to update ensembles with multiple correlated elements based on new information is shown in Figure 3-2. First, the exploration drill information with multiple elements is de-correlated using minimum/maximum autocorrelation factors (MAF) (Desbarats and Dimitrakopoulos, 2000). The de-correlated MAF factors are then used to generate initial ensembles. The new information acquired in the mining complex about the quality of the materials is de-correlated using MAF. Then, the new decorrelated information and the initial ensembles are used in the EnKF method to generate the updated ensembles of multiple correlated elements. The updated ensembles are finally transformed back from MAF

113

factors into correlated elements and averaged to mining block sizes that represent the selectivity of the operation in the mining complex.

3.2.1.1 *Updating algorithm*

A mineral deposit is discretized into an array of three-dimensional volumes referred to as mining blocks. The mining blocks are further discretized into multiple internal nodes. Let $\mathbb{Z}_e^{t',s}(x), \forall s \in S$, be a realization of the vector of the spatial random field consisting of elements $\mathbb{z}_e^{t',s}(x_i)$. $\mathbb{z}_e^{t',s}(x_i)$ represents the simulated MAF value of element e at location x_i at time t' under the scenario s, with $i \in [1, \mathcal{N}]$ being the index of internal nodes. Initial ensembles of MAF values are represented by $\mathbb{Z}_e^{t',s}(x)$ for the multiple elements in the mineral deposit. Let matrix $A_{t'}$ describe the contribution of each internal node at the location x_i at time t' towards the new information observed in the mining complex. The new information observed at time t' is also de-correlated using MAF into MAF factor $l_e^{t'}$ for element e. The Gaussian assumption in the ensemble Kalman filter is handled by transforming $\mathbb{Z}_e^{t',s}(x), \mathbb{z}_e^{t',s}(x_i)$, and $l_e^{t'}$ using the Gaussian anamorphosis function, Φ_G^e. The transformed vectors, $U_e^{t',s}(x)$, $u_e^{t',s}(x_i) = \Phi_G^e\left(\mathbb{z}_e^{t',s}(x_i)\right)$, and $m_e^{t'} = \Phi_G^e\left(l_e^{t'}\right)$ are then used in the EnKF updating process. $U_e^{t',s}(x)$ is the vector of elements $u_e^{t',s}(x_i)$. A random noise $\epsilon_e^{t'}$ is added in the new information to represent the noise with the measurement of new information as shown in Eq. 3.1. The model-based prediction $P_e^{t',s}$, which represents the predictions based on initial ensembles at the location of observed information is calculated by Eq. 3.2.

$$o_e^{t'} = m_e^{t'} + \epsilon_e^{t'} \qquad\qquad , \forall e \in E \qquad\qquad (3.1)$$

$$P_e^{t',s} = A_{t'} \cdot U_e^{t',s}(x) \qquad\qquad , \forall e \in E, s \in S \qquad\qquad (3.2)$$

$$U_e^{t'+1,s}(x) = U_e^{t',s}(x) + K_e^{t'} \cdot \left(o_e^{t'} - P_e^{t',s}\right) \qquad , \forall e \in E, s \in S \qquad\qquad (3.3)$$

$$K_e^{t'} = (A_{t'}^T \cdot C_{u_e u_e}^{t'} \cdot A_{t'} + C_{o_e o_e}^{t'})^{-1} A_{t'}^T \cdot C_{u_e u_e}^{t'} \qquad , \forall e \in E \qquad\qquad (3.4)$$

EnKF uses Eq. 3.3 to update the initial ensembles $U_e^{t',s}(x)$ with the new information based on the difference between new information and model-based predictions, and the Kalman gain. The Kalman gain, $K_e^{t'}$, is calculated using Eq. 3.4 and defines the significance of the model compared to the new information through the error covariance matrix of the model, $C_{u_e u_e}^{t'}$, and observations, $C_{o_e o_e}^{t'}$. For instance, if the new information is inaccurate, then the term $C_{o_e o_e}^{t'}$, will be high, which results in low Kalman gain.

Figure 3-2 Updating stochastic simulations of mineral deposits with new information

A low value of Kalman gain indicates a noisy observation and, therefore, the initial ensembles are not updated. On the other hand, if the Kalman gain is large, meaning the new information is accurate, then the initial ensembles are updated with the new information.

$$C_{u_e u_e}^{t'}(x_i) \cong \frac{1}{|S|} \sum_{s=1}^{S} \left(u_e^{t,s}(x_i) - \overline{u_e^{t,s}(x_\iota)} \right) \cdot \left(u_e^{t,s}(x_i) - \overline{u_e^{t,s}(x_\iota)} \right)^T , \forall i \in [1, \mathcal{N}], e$$

$$\in E \tag{3.5}$$

EnKF approximates the model error covariance matrix with a finite set of ensembles as shown in Eq. 3.5. The measurement error covariance matrix, $C_{o_e o_e}^{t'}$, is initialized randomly from a standard normal distribution. The updated ensemble values are back-transformed using Gaussian inverse transformation function, $\Phi_G^{e-1}\left(U_e^{t'+1,s}(x) \right)$, to generate updated MAF ensemble values $\mathbb{Z}_e^{t'+1,s}(x)$. The updated MAF ensemble values are further back-transformed using the MAF inverse transformation function and averaged to generate values of different elements in the mining blocks for different ensembles using Eq. 3.6.

$$d_e^{t'+1,s}(b) \approx \frac{1}{V} \sum_{i=1}^{V} \Phi_M^{e-1}\left(\mathbb{z}_e^{t'+1,s}(x_i) \right) , \forall x_i \in b, b \in B, s \in S, e \in E \tag{3.6}$$

3.2.2 Updating short-term destination policies in a mining complex

The method proposed to update the short-term destination policies of materials in a multiple product mining complex uses policy gradient reinforcement learning with neural network agents and extends upon the work of Paduraru and Dimitrakopoulos (2019). The method accounts for the uncertainty in the supply of different materials and the performance of equipment. A short-term stochastic model detailed in Sect.

3.2.2.1 is used in the policy gradient reinforcement learning framework presented in Sect. 3.2.2.2 to train the neural network destination policies.

3.2.2.1 *A stochastic model of a mining complex*

A stochastic model of a mining complex is presented in this section that uses concepts from discrete event simulation, stochastic modelling, and system dynamics to calculate the total time to move materials out of the mineral deposits. Consider an illustrative example shown in Figure 3-3, where the materials are first loaded into trucks at mine m, with shovels, that have an uncertain performance with regards to productivity, breakdown time, and repair time. Uncertainty scenarios for the shovel performance are generated from historical data. The loaded materials in the trucks are then hauled to different destinations. The decision of hauling the materials to a destination is based on destination policies, which, in this work, are neural networks that are trained through policy gradient reinforcement learning. Uncertainty scenarios for truck performance (cycle time) are also generated from historical data. Depending on the performance of the destinations, the trucks at different destinations might have a waiting time. The total extraction time $T^E_{m,d,s}$ to mine materials from mine m until it is processed at destination d under joint uncertainty scenario s, is therefore, a function of loading time $T^l_{m,s}$, hauling time to a destination $T^h_{m,d,s}$, and wait time at a destination $T^q_{d,s}$, and is calculated using Eq. 3.7.

$$T^E_{m,d,s} = f\left(T^l_{m,s}, T^h_{m,d,s}, T^q_{d,s}\right) \quad , \forall m \in \mathcal{M}, d \in \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{W}, s \in \mathbb{S} \quad (3.7)$$

Here, $\mathcal{C}$, $\mathcal{L}_S$, and $\mathcal{W}$ define the set of crushers, sulphide leach pads, and waste dumps in the mining complex. The trucks return to the same shovel they came from after delivering the materials to a destination. The neural network does not decide

117

which shovel the truck should go after it has finished delivering materials. The materials are crushed at the crushers and then conveyed to one of the processing mills with the highest available capacity (processing stream utilization). The processing mills recover the metal from the materials and generate multiple products in the mining complex. The recovery of the processing mills is also uncertain and depends on the quality of the feed materials. The stochastic scenarios of equipment performance and processing mills recovery are combined with the stochastic simulations of mineral deposits to generate the joint uncertainty scenarios $\mathbb{S}$. For instance, 15 orebody and 15 equipment performance scenarios will result in 225 joint uncertainty scenarios.

Figure 3-3 An illustrative example of a stochastic model of a mining complex

3.2.2.2 *Updating algorithm*

The stochastic model of a mining complex presented in Sect. 3.2.2.1 simulates the flow of materials in the mining complex under the joint uncertainty scenario $\mathbb{S}$, which is used to train the neural network destination policies. Note, the proposed

model decides the destination of materials based on multiple elements in a mining complex, given a fixed extraction sequence.

Figure 3-4 Process of training the neural network and adapting to new information

The complete training process of the neural network is presented in Figure 3-4(a). The joint uncertainty scenarios are fed to the stochastic model to perform the extraction and hauling simulations that generate information about the input state (SV_i), which includes the quality and quantity of materials extracted, hauled, crushed, leached, and discarded under joint uncertainty scenarios. SV_i is fed to input neurons in the fully connected feed-forward neural network. The input to different hidden neurons (h_j) is calculated using Eq. 3.8. Equation 3.9 is used to calculate the output of hidden neurons using the rectified linear function (Nair and Hinton, 2010). The input to output neurons (o_k) is then calculated using Eq. 3.10. The weights w_{ij}^h and w_{jk}^o represent the weight associated with arcs from input (i) to hidden (j) and hidden to output (k) neurons.

$$input\left(h_j\right) = \sum_{i \in nI} w_{ij}^h SV_i \qquad\qquad , \forall\, j \in nH \qquad\qquad (3.8)$$

$$output\left(h_j\right) = \max\left(0, input\left(h_j\right)\right) \qquad\qquad , \forall j \in nH \qquad\qquad (3.9)$$

$$input(o_k) = \sum_{j \in nH} w_{jk}^o \cdot output\left(h_j\right) \qquad\qquad , \forall k \in nO \qquad\qquad (3.10)$$

$$z_{b,d,t} = \frac{e^{input(o_k)}}{\sum_k e^{input(o_k)}} \qquad\qquad , \forall t \in \mathbb{T}, b \in \mathcal{B}_m, d \in \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{W} \quad (3.11)$$

The output from output neurons generates the probabilities for different actions $z_{b,d,t}$, that determine if (1) or not (0) a block b is sent to a destination d in a period t, as shown in Eq. 3.11. The probabilities are sampled to select an action during the training phase of the proposed algorithm. Sampling the probabilities in Eq. 3.11 also ensures that the blocks are only assigned to one destination. Equations 3.12 and 3.13 are then used to calculate the amount of metal property a, and mass respectively at the different destinations i.

$$v_{a,i,t,s} = \sum_{b \in \mathcal{B}_m} g_{a,b,s} \cdot m_{b,s} \cdot z_{b,d,t} \quad , \forall t \in \mathbb{T}, a \in \mathbb{P}_M, i \in \mathcal{C} \cup \mathcal{L}_S, s \in \mathbb{S} \quad (3.12)$$

$$v_{a,i,t,s} = \sum_{b \in \mathcal{B}_m} m_{b,s} \cdot z_{b,d,t} \quad , \forall t \in \mathbb{T}, a \in \mathbb{P}_T, i \in \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{W}, s \in \mathbb{S} \quad (3.13)$$

The materials from the different destination $i \in \mathcal{C} \cup \mathcal{L}_S$ is further sent to different processing streams $j \in \mathcal{P} \cup \mathcal{L}_O$. Processing stream utilization decisions $y_{a,i,j,t,s}$, represents the amount of materials property a, sent from destination i to j in period t, under scenario s, and is decided based on available capacity at the different processing streams. Equation 3.14 is used to calculate the materials at the different processing

streams in the mining complex. Equation 3.15 ensures that flow conservation is preserved with the processing stream utilization decisions.

$$v_{a,j,t,s} = \sum_{i \in \mathcal{C}} y_{a,i,j,t,s} \cdot v_{a,i,t,s} \quad , \forall t \in \mathbb{T}, a \in \mathbb{P}_M \cup \mathbb{P}_T, j \in \mathcal{P} \cup \mathcal{L}_O, s \in \mathbb{S} \quad (3.14)$$

$$\sum_{j \in \mathcal{P} \cup \mathcal{L}_O} y_{a,i,j,t,s} = 1 \quad , \forall t \in \mathbb{T}, i \in \mathcal{C}, s \in \mathbb{S} \quad (3.15)$$

$$v_{a,i,t,s} - d^+_{a,i,t,s} \le U_{a,i,t} \quad , \forall \ t \in \mathbb{T}, a \in \mathbb{P}_M, i \in \mathcal{P} \cup \mathcal{L}_S \cup \mathcal{L}_O, s \in \mathbb{S} \quad (3.16)$$

$$v_{a,i,t,s} + d^-_{a,i,t,s} \ge L_{a,i,t} \quad , \forall \ t \in \mathbb{T}, a \in \mathbb{P}_M, i \in \mathcal{P} \cup \mathcal{L}_S \cup \mathcal{L}_O, s \in \mathbb{S} \quad (3.17)$$

$$f(X) = \underbrace{\frac{1}{|\mathbb{S}|} \sum_{s \in \mathbb{S}} \sum_{t \in \mathbb{T}} \sum_{i \in \mathcal{P} \cup \mathcal{L}_O \cup \mathcal{L}_S} \sum_{a \in \mathbb{P}_M} P_{a,i} \cdot v_{a,i,t,s} \cdot r_{a,i,s}}_{\text{Part I}}$$

$$\underbrace{-\frac{1}{|\mathbb{S}|} \sum_{s \in \mathbb{S}} \sum_{t \in \mathbb{T}} - \sum_{i \in \mathcal{P} \cup \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{L}_O \cup \mathbb{M}} \sum_{a \in \mathbb{P}_T} C_{a,i} \cdot v_{a,i,t,s}}_{\text{Part II}}$$

$$\underbrace{-\frac{1}{|\mathbb{S}|} \sum_{s \in \mathbb{S}} \sum_{t \in \mathbb{T}} \sum_{i \in \mathcal{P} \cup \mathcal{L}_S \cup \mathcal{L}_O} \sum_{a \in \mathbb{P}_M} \left(c^+_{a,i} \cdot d^+_{a,i,t,s} + c^-_{a,i} \cdot d^-_{a,i,t,s} \right)}_{\text{Part III}} \quad (3.18)$$

Equations 3.16 and 3.17 are used to calculate the amount of deviation from different production targets in the mining complex. The metal is finally recovered at the different processing destinations. The objective/cash flow/reward function is given by Eq. 3.18. Part I in the objective function represents the profits from selling different products; Part II represents the different costs incurred throughout the flow

of materials, and Part III represents the penalties incurred due to deviation from different production targets. The objective function is an expected value. Equations 3.12-3.18 are based on recent developments in stochastic mine planning models (Montiel and Dimitrakopoulos, 2015; Goodfellow and Dimitrakopoulos, 2016; Quigley and Dimitrakopoulos, 2019). Policy gradient reinforcement learning (Sutton et al., 2000) offers the ability that, given a reward function f and probability density function z_W parameterized by W, the equality in Eq. 3.19 below holds true.

$$\nabla_W E_{x \sim z_W(x)}[f(x)] = E_{x \sim z_W(x)}\left[f(x)\nabla_W \log\left(z_W(x)\right)\right] \tag{3.19}$$

In Eq. 3.19, $f(x)$ corresponds to the reward function calculated using Eq. 3.18 and $z_W(x)$ corresponds to the action-selection probabilities computed using Eq. 3.11. The weight matrix, W, contains the values of the hidden w_{ij}^h and the output neurons w_{jk}^o. As it is common in stochastic gradient methods (Bottou, 2010), $E_{x \sim z_W(x)}\left[f(x)\nabla_W \log\left(z_W(x)\right)\right]$ is replaced with $f(X)\nabla_W \log\left(z_W(X)\right)$, where $f(X)$ represent the cumulative reward obtained during the planning horizon $\mathbb{T}$ using the vector of decisions X. The gradient of $\log\left(z_W(X)\right)$ can, therefore, be calculated using Eq. 3.20, where the sum is over the planning horizon and over the destinations. Finally, the stochastic approximation of $\nabla_W E_{x \sim z_W(x)}[f(x)]$ can be computed using Eqs. 3.18-3.20.

$$\nabla_W \log\left(z_W(X)\right) = \sum_{t \in \mathbb{T}} \sum_{d \in \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{W}} \nabla_W \log z_W(d)\, z_{b,d,t} \tag{3.20}$$

$$g_{i+1} = \gamma g_i + (1-\gamma)\nabla_W E_{x \sim z_W(x)}[f(x)]^2 \qquad , \forall\, i \in [1, nIter] \tag{3.21}$$

$$W_{i+1} = W_i + \frac{\eta \nabla_W E_{x \sim z_W(x)}[f(x)]}{\sqrt{g_{i+1}} + \partial} \qquad , \forall i \in [1, nIter] \qquad (3.22)$$

The weight matrix $W = \{w_{ij}^h, w_{jk}^o\}$ of the neurons in the neural network is initialized randomly and updated using the gradient ascent method named RMSprop (Hinton et al., 2012). The RMSprop method uses Eqs. 3.21 and 3.22 to backpropagate and update the weight of the neurons in the training phase of the neural network. This process (Eqs. 3.8-3.22) continues, and the neural network is trained until the pre-defined stopping criteria, $nIter$, is reached.

The training phase of the neural network allows the generation of destination policies that can adapt to new information. Figure 3-4(b) represents the process of adapting the neural network destination policies when new information is acquired in a mining complex. The new information is first used to update the joint uncertainty scenarios using the method outlined in Sect. 3.2.1. The updated joint uncertainty scenarios are then fed to the stochastic model outlined in Sect. 3.2.2, which simulates the extraction and hauling of materials. The information from the previous step is fed to the trained neural network that decides the destination of materials (the action with maximum probability is selected) and the materials from such destinations are then sent to one of the processing streams based on the available capacity of the different processing streams. Finally, the forecasts for the different production targets are calculated using Eqs. 3.12-3.18 and further evaluated regarding their probability of meeting the different production targets. The neural network is retrained for a few iterations if the production targets are not met to adjust the weight of the neural network and better meet the production targets.

123

3.3 Application at a copper mining complex

The proposed framework for updating the short-term destination of materials is applied at a copper mining complex, which demonstrates the applied aspects of the proposed method. In the case study, the blasthole data collected during the mine's operation is used to update the stochastic simulations of mineral deposits with multiple elements. The neural network destination policies account for uncertainty in (a) supply of multiple materials with multiple elements, (b) performance of equipment related to its availability, cycle times, utilization, downtime, repair time, and productivity, and (c) recovery of metal in processing mills. However, the framework is flexible to include different types of new information in the updating framework. The implementation assumes that the mining complex has the necessary infrastructure related to wireless internet server/system and cloud services to handle, store, and transmit the new collected information and feedback the adapted short-term production plan to the mining operation, as it is the case in mining complex involved in the application present herein.

3.3.1 Overview of the copper mining complex

The copper mining complex consists of two mineral deposits (A and B) with mining blocks of size 25x25x15 m^3. The mineralization has eight different mine zones each. The materials are extracted from both deposits and are sent to one of the seven destinations (five crushers, one sulphide leach pad, and one waste dump), as shown in Figure 3-5. For measuring the performance of the proposed framework, a part of the deposit that consists of 5,581 mining blocks in each deposit extracted over

210 days is used. Materials from five different crushers are then processed at three different processing mills and an oxide leach pad.

Figure 3-5 The copper mining complex

The materials from the leach pads are sent to a copper cathode plant that produces copper cathodes. The processing mills generate copper concentrate as the primary product and gold (Au), silver (Ag), and molybdenum (Mo) concentrate as secondary products, which are transported to the port. The products from the port and copper cathode plant are finally transported and sold to different customers and/or the spot market. Additional details about the case study are presented in Appendix 3.2.

3.3.2   Cut-off grade vs adaptive neural network destination policies

The copper mining complex currently uses a single element (copper) predefined cut-off grade based destination policies optimized using Lane's theory (Lane, 1984, 1988; Rendu, 2014) and cannot account for new information collected during the mine's operation. The copper mining complex is a major producer of copper products

125

and does not consider secondary products in the optimization of its cut-off grade destination policies. The details of the cut-off grade destination policies are outlined in Table 3-1. First, the materials are classified as sulphide high grade (SHG), sulphide low grade (SLG), oxide based on the materials classification criteria (i.e., ratio of soluble copper (CuS) to total copper (CuT)). The materials classification criteria are necessary to determine the possible processing destinations allowed to process the materials. The cut-off grade destination policies then use the cut-off grades specified in Table 3-1 to determine the destination at which the material will be processed.

Table 3-1 Material classification criteria and cut-off grade destination policies used at the copper mining complex

Materials classification	Materials classification criteria	Cut-off grade destination policies	Destination
SHG	$\dfrac{CuS}{CuT} \leq 0.2$	$CuT \geq 0.6$	Processing Mill
		$0.3 \leq CuT < 0.6$	Sulphide Leach Pad
		$CuT < 0.3$	Waste Dump
SLG	$0.2 < \dfrac{CuS}{CuT} < 0.5$	$CuT > 0.3$	Sulphide Leach Pad
		$CuT \leq 0.3$	Waste Dump
Oxide	$\dfrac{CuS}{CuT} \geq 0.5$	$CuS \geq 0.2$	Oxide Leach Pad
		$CuS < 0.2$	Waste Dump

The neural network destination policies decide the destination of mining blocks based on the properties of multiple elements in a mining block, as well as the performance of and interaction between the different components of the mining complex. In addition, the proposed method adapts such destination decisions of mining blocks with incoming new information in the mining complex (See Sect. 3.2.2). Similar to the cut-off grade destination policies, the materials are first characterized as SHG, SLG, and oxide, based on the material classification criteria

mentioned in Table 3-1 to find the allowed processing destinations for a mining block. However, instead of using the cut-off grade destination policies mentioned in Table 3-1, the neural network destination policies are used to decide the destination of such materials. As mentioned in Sect. 3.2.2, the neural networks decide whether (1) or not (0) to process the materials at (i) the processing mills, (ii) a sulphide leach pad, or (iii) an oxide leach pad. Four different neural networks are built and trained using policy gradient reinforcement learning. The different neural networks are designed to decide the destinations of extracted materials, which can be either sulphides or oxides. The first neural network is for SHG materials extracted from mine A, the second is for SHGs from mine B, the third is for SLGs from mines A and B, and the fourth is for oxides from mines A and B. The first neural network takes 32 inputs, has 800 hidden neurons, and generates 5 outputs. The five outputs each correspond to the probability of selecting crusher 1, crusher 2, crusher 3, the sulphide leach pad, and the waste dump, as the destination of the materials. The second neural network takes 20 inputs, has 600 hidden neurons, and generates 3 outputs. The three outputs correspond to the probability of selecting crusher 5, the sulphide leach pad, and the waste dump as the destination of materials. The third and fourth neural networks take 7 inputs, have 300 hidden neurons, and generate 2 outputs. The two outputs from third neural network correspond to the probability of selecting the sulphide leach pad and waste dump as the destination of materials. The two outputs from the fourth neural network correspond to the probability of selecting crusher 4 (oxide leach crusher) and the waste dump as the destination of materials.

3.3.3 Parameter selection

This section discusses the selection of different parameters associated with the proposed adaptive neural network destination policies. The state vector information SV_i consists of 7 to 32 different types of information depending on the complexity of the processing destination and are fed to the input neurons of the neural network. For instance, the SV_i for the processing mill neural network consists of information about the mass of a mining block, different elements such as total copper, soluble copper, arsenic, gold, silver, and molybdenum in the mining block, the materials being crushed and leached, the performance of equipment, and the wait times at the crushers. Similarly, the number of hidden neurons in the neural network ranges from 300 to 800, depending on the number of input neurons. There are only two output neurons to decide whether (1) or not (0) the mining block is processed at the respective destination. The learning rate and the decay rate with the neural network is set to 10^{-3} and 0.99, respectively, as suggested in Hinton et al., (2012). The smoothing term is set to 10^{-6} (Ruder, 2016). The weight of the neurons in the neural network is initialized randomly using the Xavier initialization (Glorot and Bengio, 2010). The number of iterations required to train the neural network is set to 7,500. The number of mineral deposit simulations to use for training the neural network is set to 15 based on the tests show in the supplementary material. The details of parameter selection are presented in Appendix 3.3.

3.3.4 Results

The results of the proposed adaptive neural network destination policies to update the short-term destination decisions with new information are presented in this

section. Results are reported using the 10th, 50th, and 90th percentile risk profiles (P10, P50, and P90 respectively) of the different performance indicators considering 100 joint uncertainty scenarios (10 equipment performance and 10 orebody scenarios). The results reported in this section are based on a set of 100 joint uncertainty scenarios that were not used to train the neural network destination policies. Testing the neural network destination policies on an unseen set of joint uncertainty scenarios shows the reliability of the proposed framework and highlights the overfitting issues, if any, with the neural network destination policies. The forecasts of the production targets with the proposed framework are compared to the forecasts of the cut-off grade destination policies over the same 100 joint uncertainty scenarios throughout its presentation and discussion to highlight the differences and added value of the adaptive framework, where appropriate. The training phase of the neural network takes about 52 hours, with 12,500 iterations on an Intel processor core i7 with 8GB of RAM. However, it only takes about five minutes to update the stochastic simulations of the two mineral deposits and to adapt the destination decisions of mining blocks for 210 days using the proposed adaptive framework. The results are presented for both the destination policies for initial and update stochastic simulations of mineral deposits. The results presented for metal production and cash flows are scaled for confidentiality purposes (mine's cut-off grade destination policies for initial simulations being 100%). Additional results from the case study are presented in Appendix 3.4 and Appendix 3.5.

3.3.4.1 *Updated stochastic simulations of mineral deposit*

Figure 3-6 shows one of the initial and updated simulations of the total copper mineral attribute of the mineral deposit A at block support. The initial stochastic simulations of six correlated elements in the two mineral deposits, conditional to the exploration drillholes' samples, are generated using a generalized sequential Gaussian simulation (Dimitrakopoulos and Luo, 2004).

Figure 3-6 Updated block simulations compared to initial block simulations for bench 1 for the mineral deposit A

Six different correlated elements: soluble copper, total copper, arsenic, gold, silver, and molybdenum, in the two mineral deposits are updated using the method discussed in Sect. 3.2.1 with the new blasthole data collected during the short-term operations. The blasthole data in a mine zone are only considered to update the mining blocks in the same mine zone to respect the geological features of the mineral deposit. It is clear from Figure 3-6 that the updated simulations maintain the significant structures inferred from the exploration drillholes data and updates the local

130

characteristics with the new blasthole data. A histogram of the initial and updated simulations at point support confirms such results, where the distribution of total copper in bench 1 for mineral deposit A is very different for the initial and updated simulations. The updated simulations show a higher proportion of high-grade copper materials, as compared to the initial simulations.

3.3.4.2 *Production target*

The forecasts for the different production targets are shown in this section for the neural network destination policies and are compared to the cut-off grade destination policies.

Figure 3-7 Forecasts of the capacity target of mill-2 with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-7(b) shows the risk profile of meeting the capacity target with mill-2 for initial simulations using neural network destination policies compared to the cut-off

grade destination policies in Figure 3-7(a). The neural network destination policies are better at meeting the target with maximum utilization of the mill's capacity, as compared to high fluctuations and lower chances of meeting the target in the cut-off grade destination policies. The neural network destination policies (Figure 3-7(d)) has increased the chance of meeting production targets compared to the high fluctuations in the cut-off grade destination policies (Figure 3-7(c)) over the updated simulations.

Figure 3-8 Forecasts of arsenic blending target of mill-2 with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-8(b) and Figure 3-8(a) show the risk of meeting the blending target of arsenic at mill-2 for initial simulations with neural network and cut-off grade destination policies, respectively. The neural network destination policies have higher chances of meeting such a target with minimal deviations only after 80 days, as

compared to the cut-off grade destination policies, which have a higher chance of deviating from such targets, more specifically during the first 80 days. The two destination policies are unable to meet the blending restrictions as shown in Figure 3-8(d) and Figure 3-8(c) over the updated simulations. The lower chances of meeting the arsenic target with the updated destination decisions are due to the fixed extraction sequence decision in the proposed framework. Therefore, if there is a high concentration of arsenic in the updated simulations, it is hard to control the arsenic concertation in the mill without adapting the extraction sequence.

3.3.4.3 *Metal production*

Figure 3-9(b) and Figure 3-9(a) represent the risk profile of cumulative copper production at the mills for the initial simulations with neural network and cut-off grade destination policies, respectively. The neural network destination policies recover 11% additional copper metal, as compared to the mine's cut-off grade destination policies for the initial simulations. The neural network destination policies recover an additional 19 % copper metal (Figure 3-9(d)), as compared to an additional 8% copper metal in the mine's cut-off grade destination policies (Figure 3-9(c)) over the updated simulations.

Figure 3-9 Forecasts of total copper production at the processing mills with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-10 Forecasts of total gold production at the processing mills with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

134

Figure 3-10 shows the risk profiles of the production of secondary product gold concentrate using the neural network and the cut-off grade destination policies. The neural network destination policies generate 27% additional gold product (Figure 3-10(b)), as compared to the mine's cut-off grade destination policies (Figure 3-10(a)) over the initial simulations. The adapted decisions of neural network destination policies generate an additional 53% of the gold product (Figure 3-10(d)), as compared to an additional 38% for the mine's cut-off grade destination policies (Figure 3-10(c)) over the updated simulations.

3.3.4.4 *Cash flows*

Figure 3-11 shows the risk profile of cumulative cash flows with the neural network and cut-off grade destination policies. The neural network destination policies present a 15% higher cumulative cash flows compared to the mine's cut-off grade destination policies for the initial simulations (Figure 3-11(a)).

Figure 3-11 Forecasts of the cumulative cash flow of the mining complex with the (a) initial cut-off grade and neural network block destinations, and (b) updated cut-off grade and neural network block destinations

The neural network destination policies generate an additional 22% cumulative cash flows, as compared to an additional 11% for the mine's cut-off grade destination policies (Figure 3-11(b)) over the updated simulations.

3.3.4.5 *Updated destination decisions*

Figure 3-12(b) shows the destination decisions of the neural network destination policies compared to the cut-off grade destination policies in Figure 3-12(a) for initial simulations. The adapted destination decisions of the neural network and the cut-off grade destination policies are shown in Figure 3-12(d) and Figure 3-12(c), respectively. The neural network destination decisions are very different from the cut-off grade destination decisions for initial and update simulations, which result in better chances of meeting production targets, consistently higher cumulative cash flows, and increased metal production.

Figure 3-12 Destination decisions of mining blocks for bench 1 in mineral deposit A with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

136

The reason for the better performance of neural network destination policies is due to its ability to:

1. Acount for and capitalize on the performance of and interaction amongst the different components in the mining complex, thus enabling complex decision-making under different sources of uncertainties.
2. Integrate multiple sources of uncertainty, such as the supply of materials, the performance of equipment, and the recovery of metal during the decision-making process
3. Account for multiple products, such as copper, gold, silver, and molybdenum, as well as deleterious elements such as arsenic, while deciding the destination of mining blocks.

3.4   Conclusions

This paper presents a novel continuous updating framework for adapting the short-term flow of materials in a mining complex with incoming new information. The framework consists of two parts: first updating uncertainty models with a new extension of ensemble Kalman filter and second, feeding the updated uncertainty models to a neural network agent (trained using policy gradient reinforcement learning) that adapts the destination decisions of extracted material. The proposed framework is applied at a copper mining complex, which shows its applied aspects and an excellent performance to respond and integrate the incoming new information efficiently in an operational mining environment for adapting the materials flow. The proposed framework better meets the capacity and blending requirements of the different processing mills of the copper mining complex compared to the mine's cut-off grade destination policies. The proposed framework generates an additional 11%,

27%, 29%, and 29% of copper, gold, silver, and molybdenum products, respectively, and an additional 15% of cash flows, as compared to the mine's cut-off grade destination policies for the initial simulation. The extended ensemble Kalman filter updates multivariate local features of the mineral deposits with new blasthole information. The neural network destination policies are better at responding to the new information and adapt the destination decisions over the updated simulations more intelligently to meet the targets better. The updated destination decisions from neural network destination policies generate an additional 19%, 53%, 71%, and 76% of copper, gold, silver, and molybdenum products, respectively, as well as an additional 22% of cash flows. The mine's cut-off grade destination policies only generate an additional 8%, 38%, 56%, and 61% of copper, gold, silver, and molybdenum products, respectively, and an additional 11% of cash flows, over the updated simulations. The proposed framework only adapts the destination decisions of the mining blocks, thus limiting the full potential and use of new information. In the future, a framework that can adapt all the relevant decisions of the short-term production plan will be developed.

Appendix 3.1

Table 3-2 outlines the notations, sets, indices, parameters, and constants used in the proposed framework. Table 3-3 shows the variables used in the proposed framework.

Table 3-2 Sets, indices, parameters, and constants used in the proposed framework

Parameters	Definition
S	Set of stochastic orebody simulations, $s \in S$
$\mathbb{S}$	Set of joint uncertainty scenarios that include orebody and equipment simulations, $s \in \mathbb{S}$
$\mathbb{T}$	Production planning horizon, $t \in \mathbb{T}$
$\mathcal{M}$	Set of mines in a mining complex, $m \in \mathcal{M}$
$\mathcal{B}_m$	Set of mining blocks in a mine $m, b \in \mathcal{B}_m$
$\mathcal{N}$	Set of internal nodes in a mine
x_i	Location of internal nodes in a mining block $b, i \in [1, \mathcal{N}]$
$\mathcal{C}$	Set of crushers in a mining complex
$\mathcal{P}$	Set of processing mills in a mining complex
$\mathcal{L}_O$	Set of oxide leach pads in a mining complex
$\mathcal{L}_S$	Set of sulphide leach pads in a mining complex
$\mathcal{W}$	Set of waste dumps in a mining complex
t'	Time step when new information is collected
V	Set of internal nodes in a mining block b
E	Set of elements in a mineral deposit, $e \in E$
$d_e^{t',s}(x_i)$	Initial data value of the internal node at the location x_i for element e at time t', and scenario s
$d_e^{t',s}(b)$	Initial data value of mining block b for element e at time t' and scenario s
$\mathbb{z}_e^{t',s}(x_i)$	Initial MAF value at location x_i for element e at time t' and scenario s
$\mathbb{Z}_e^{t',s}(x)$	Vector of $\mathbb{z}_e^{t',s}(x_i)$ for element e at time t' and scenario s
$A_{t'}$	Matrix of the contribution of internal nodes towards new information at time t'
$A_{t'}^T$	Transpose of matrix $A_{t'}$
$\epsilon_e^{t'}$	Error in the new information for element e at time t'
$P_e^{t',s}$	Model-based prediction for element e at time t' and scenario s

$K_e^{t'}$	Kalman gain for element e at time t'
$C_{o_e o_e}^{t'}$	Measurement error covariance matrix for element e at time t'
$C_{u_e u_e}^{t'}$	Model error covariance matrix for element e at time t'
$u_e^{t',s}(x_i)$	Initial Gaussian values for element e at time t'generated by transforming MAF values at the location x_i
$U_e^{t',s}(x)$	Vector of $u_e^{t',s}(x_i)$ for elements e at time t'
$m_e^{t'}$	MAF value of new information transformed to Gaussian values for element e at time t'
$U_{a,i,t}$	Upper production limit for property a at i in period t
$L_{a,i,t}$	Lower production limit for property a at i in period t
$C_{a,i}$	Cost of processing material property a at i
$c_{a,i}^+$	Cost of deviation from the upper target $U_{a,i,t}$ for material property a at i
$c_{a,i}^-$	Cost of deviation from the lower target $L_{a,i,t}$ for material property a at i
nI	Number of input neurons
nH	Number of hidden neurons
nO	Number of output neurons
h_j	Hidden neuron j
o_k	Output neuron k
$m_{b,s}$	Mass of block b under scenario s
$g_{a,b,s}$	Grade of material property a in block b under scenario s
$\epsilon_e^{t'}$	Noise in the new information for element e at time t'
$nIter$	Number of training iterations
g_i	Gradient at iteration i, $i \in [1, nIter]$
η	Decay rate
∂	Smoothing term
γ	Learning rate
Φ_M^e	MAF transformation function for element e
Φ_G^e	Gaussian transformation function for element e
Φ_M^{e-1}	MAF inverse transformation function for element e
Φ_G^{e-1}	Gaussian inverse transformation function for element e
$C_{u_e u_e}^{t'}$	Model error covariance matrix for element e at time t'
$C_{o_e o_e}^{t'}$	Measurement error covariance matrix for element e at time t'
$l_e^{t'}$	MAF value of new information collected for element e at time t'
$\mathbb{P}_T$	Property tonnage that flows in the mining complex

Variables	Definition
$\mathbb{P}_M$	Set of metal properties that flow in the mining complex
$T^l_{m,s}$	Loading time with the shovel at mine m under joint uncertainty scenario s

Table 3-3 Variables used in the proposed framework

Variables	Definition
$z_{b,d,t} \in \{0,1\}$	Defines if (1) or not (0) a block b is sent to destination d in period t
$y_{a,i,j,t,s} \in [0,1]$	Amount of property a send from i to j in period t under joint uncertainty scenario s
$d^+_{a,i,t,s} \in \mathbb{R}$	Excess from target $U_{a,i,t}$ for material property a at i in period t under joint uncertainty scenario s
$d^-_{a,i,t,s} \in \mathbb{R}$	Shortage from target $L_{a,i,t}$ for material property a at i in period t under joint uncertainty scenario s
$v_{a,i,t,s} \in \mathbb{R}$	Amount of material property a at i in period t under joint uncertainty scenario s
$r_{a,i,s} \in [0,1]$	Recovery of material property a at i under joint uncertainty scenario s
$U_e^{t'+1,s}(x)$	Updated Gaussian values
$\mathbb{z}_e^{t'+1,s}(x_i)$	Updated MAF value at location x_i for element e at time $t'+1$ and scenario s
$\mathbb{Z}_e^{t'+1,s}(x)$	Vector of $\mathbb{z}_e^{t'+1,s}(x_i)$ for element e at time $t'+1$ and scenario s
$d_e^{t'+1,s}(b)$	Updated data value of mining block b for element e at time t' and scenario s
$P_{a,i}$	Profit of product a at location i. The profit is calculated after deducting all the costs incurred to generate the products
$T^h_{m,d,s}$	Hauling time from mine m to destination d under joint uncertainty scenario s
$T^q_{d,s}$	Queue time at destination d under joint uncertainty scenario s
$T^E_{m,d,s}$	Total extraction time from mine m to destination d under joint uncertainty scenario s
SV_i	Components of input state vector fed to input neuron i
w^h_{ij}	Weight with an arc connecting input neuron i to hidden neuron j
w^o_{jk}	Weight with an arc connecting hidden neuron j to output neuron k

Appendix 3.2

This section outlines the additional details of the case study presented in the paper. The production targets with different components of the mining complex are presented in Table 3-4. The operational and economic parameters used in the case study are outlined in Table 3-5. The production targets and economic parameters are scaled for confidentiality purposes.

Table 3-4 Production target for copper mining complex

Attribute	Value
Crusher 1, 2, 3, 4, and 5 (Per Day)	13.19, 25.69, 32.06, 20.83, and 29.06 tonne
Mill 1, 2, and 3 (Per Day)	28.4, 34.02, and 38.88 tonne
Oxide and sulphide leach pad (Per Day)	20.83 and 93.7 tonnes
Arsenic grade limit (Per Day)	100 %

Table 3-5 Operational and economic parameters for copper mining complex

Attribute	Value
Number of mining blocks – mineral deposit A and B	5581 and 5581
Scheduling period	210 days
Recovery of copper at oxide and sulphide leach pad	0.651 and 0.275
Recovery of copper at processing mills	Stochastic
Recovery of gold, silver, and molybdenum at processing mills	0.25
Processing cost – mills, oxide leach pad, and sulphide leach pad	3.82, 6.4, and 1.84 \$/tonne
Selling cost – mills, oxide leach pad, and sulphide leach pad	0.26, 0.25, and 0.25 \$/lb
Selling price – copper, gold, silver, and molybdenum	5511, 35.2×10^6, 4.9×10^5, and 1.3×10^4 \$/ tonne
Fixed processing cost – Mill 1, 2, and 3	11925.5, 14252.5, and 16288.5 \$/ hour

Penalty arsenic grade (As)	3 $/ % above limit

Appendix 3.3

This section presents the parameter selection criteria of the adaptive neural network destination policies presented in the paper. A critical parameter is the number of stochastic simulations of mineral deposits and the number of iterations used to train the adaptive neural network destination policies. The basis for selecting such parameters is the evolution of objective function. The iteration number at which the objective function plateaus is selected for training the model. Figure 3-13 shows one such test with the progression of the objective function in the training phase of the neural network with 15 mineral deposits simulations (S). At around 7,500 training iterations, the objective function plateaus for 15 mineral deposits simulations, which presents the relation in Eq. 3.23.

$$nIter = 500 * S \qquad (3.23)$$

Figure 3-13 Progression of the objective function with varying iteration number

Table 3-6 Comparison of training the neural network destination policies with varying number of mineral deposits simulations

Destination (Number of blocks)	Number of Simulations (S)						
	5	10	15	20	25	30	35
Mill	2040	2008	1967	2053	2155	2070	2095
Sulphide leach pad	3111	3357	3447	3758	3303	3218	3544
Oxide leach pad	308	301	301	294	304	301	308
Waste	4010	3758	3873	3431	3515	3978	3705
Objective Function (%)	111	111	112	114	115	115	115

The number of simulations of mineral deposits to use for training the neural network is decided by training the network with different values of S and, then, testing the trained network on a new set of 100 joint uncertainty scenarios (which includes 10 orebody and 10 equipment performance scenarios) to compare the performance in terms of meeting different production targets. The value of the S at which the objective function stabilizes is selected to be used for training the neural network destination policies. The details of this study are shown in Table 3-6. From Table 3-6, it can be seen that after 15-20 simulations the number of decisions does not change significantly and also the change in the objective function value is negligible, which highlights that 15-20 simulations are enough to achieve a stable solution for the proposed short-term adaptive neural network destination policies.

Appendix 3.4

This section outlines the additional results of updating the stochastic simulations of mineral deposits with new blasthole data. Table 3-7 represents the number of blasthole data used to update the two benches in the two mineral deposits in the copper mining complex.

Figure 3-14 Updated block simulations compared to initial simulations for bench 2 for mineral deposit A

Figure 3-14 represents one of the initial and updated stochastic simulations of the copper total mineral attribute in the mineral deposit A for bench 2 at block support with the new blasthole data. The updated simulations preserve the structures inferred from the exploration drillhole samples and update the local features of the deposits with the new blasthole data. The histograms of initial and updated simulations at point support in Figure 3-14 also show that the updated simulation has a higher proportion of high-grade materials compared to initial simulations.

Table 3-7 Blasthole data used to update the stochastic simulations of the two mineral deposits

Bench number	Number of blasthole data	
	Mineral deposit A	Mineral deposit B
Bench 1	3437	222
Bench 2	3309	258

Appendix 3.5

This section outlines the additional results of updating and adapting the destination of mining blocks with the proposed framework.

Figure 3-15 Destination decisions of mining blocks for bench 2 in mineral deposit A with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-15 represents the destination decisions of mining blocks for bench 2 in the mineral deposit A using the neural network destination policies and cut-off grade destination policies. Figure 3-15(b) and Figure 3-15(a) represent the destination decisions of the neural network destination policies and cut-off grade destination policies respectively for initial simulations. The updated decisions with neural network destination policies and cut-off grade destination policies are shown in Figure 3-15(d) and Figure 3-15(c), respectively.

Figure 3-16 Forecasts of the capacity target of mill-1 with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-16 (b) and (a) shows the risk profiles of the capacity target with mill-1 using neural network and cut-off grade destination policies for initial simulations. The neural network policies meet the capacity targets better than the cut-off grade destination policies. Figure 3-16(d) and (c) shows the risk of meeting capacity target with mill-1 when the decisions are adapted with neural network and cut-off grade destination policies, respectively. It is observed that both the policies have lower chances of meeting the mill-1 target arising from combined uncertainty in the supply of materials from both the mineral deposits.

Figure 3-17 Forecasts of the capacity target of mill-3 with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-17(b) and (a) represents the risk profiles for mill-3 for both the policies for initial simulations. The neural network policies can meet the target better than the cut-off grade destination policies, which has high fluctuations and lower chances of meeting the mill-1 target. The adapted destination decisions with the neural network (Figure 3-17(d)) has increased chances of meeting the mill-1 target compared to the cut-off grade destination policies (Figure 3-17(c)).

148

Figure 3-18 Forecasts of arsenic blending target of mill-1 with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-19 Forecasts of arsenic blending target of mill-3 with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-18 and Figure 3-19 show the risk profile of arsenic blending target for mill-1 and mill-3, respectively. The destination decisions with the neural network and cut-off grade policies show very few deviations from such blending targets over the initial simulations, however, with updated simulations, such targets will be violated, and large deviations are expected with both the policies. Since the extraction sequence decisions are fixed in the proposed framework and the model is only deciding the destination of materials. Therefore, the high concentration of arsenic in the simulations cannot be controlled at the processing mills without adapting the extraction sequence.

Figure 3-20 Forecasts of total silver production at the processing mills with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

Figure 3-21 Forecasts of total molybdenum production at the processing mills with the (a) initial cut-off grade block destinations, (b) initial neural network block destinations, (c) updated cut-off grade block destinations, and (d) updated neural network block destinations

The neural network destination policies generate an additional 29% higher silver product (Figure 3-20(b)), and an additional 29% higher molybdenum product (Figure 3-21(b)), as compared to the mine's cut-off grade destination policies (Figure 3-20(a) and Figure 3-21(a)) over the initial simulations. The adapted decisions of neural network destination policies generate an additional 71% and 76% of silver and molybdenum products respectively (Figure 3-20(d) and Figure 3-21(d)) over the updated simulations. The mine's cut-off grade destination policies generate an additional 56% and 61% of silver and molybdenum products respectively (Figure 3-20(c) and Figure 3-21(c)) over the updated simulations. The updated decisions with the present mine's cut-off grade destination policies produce significantly less

quantity of secondary products compared to the updated decisions from adaptive neural network destination policies.

CHAPTER 4
Production Scheduling in Industrial Mining Complexes with Incoming New Information using Tree Search and Deep Reinforcement Learning

The previous chapter presented a self-learning artificial intelligence framework for adapting the flow of materials with incoming new information in a mining complex. This chapter develops a new self-play reinforcement learning algorithm for adapting all the major short-term production planning decisions simultaneously, i.e. adapting the short-term sequence of extraction, the destination of materials, and the utilization of processing streams simultaneously with incoming new information.

4.1   Introduction

Artificial intelligence (AI) algorithms have already been developed for applications in different engineering fields, but have not been developed and extended for short-term production scheduling in industrial mining complexes (Matamoros and Dimitrakopoulos, 2016; Quigley and Dimitrakopoulos, 2019). A mining complex is an integrated value chain with multiple interrelated components, such as raw material suppliers (mineral deposits), heavy machinery (shovels and trucks), destinations (crushers, stockpiles, and waste dumps), processing streams (processing mills and leach pads), tailings, and customers. Heavy machinery extracts and transports raw materials to the destinations. The materials from the destinations are then transported to processing streams, which process the materials to generate products that are

delivered to different customers. A long-term production schedule is developed for a mining complex to provide the annual strategic decisions, targets, and forecasts, while maximizing the cumulative discounted cash flows and accounting for supply and market uncertainties (Montiel and Dimitrakopoulos, 2015; Goodfellow and Dimitrakopoulos, 2016; Mai et al., 2019; Paithankar and Chatterjee, 2019; Paithankar et al., 2020). A short-term production schedule, i.e. an operational production schedule at a monthly/weekly/daily timescale, is then generated within the predefined long-term schedule, and aims to ensure compliance with the long-term targets while maximizing cash flows. The major short-term production scheduling decisions in a mining complex are extraction sequencing, destination policies, and processing stream utilization. Extraction sequencing refers to determining the location and time of raw material extraction from the mineral deposit while adhering to mine slope stability and equipment movement restrictions. Mine slope stability constraints state that a block cannot be extracted safely until all overlying blocks within a predefined inclination are extracted first. Equipment movement restrictions state that a block cannot be extracted until one of the surrounding blocks is extracted first to provide access. Destination policies refer to finding the destination of the extracted raw materials while satisfying material eligibility conditions. Material eligibility conditions state that specific types of materials cannot go to some destinations due to limitations with their downstream recovery process. Processing stream utilization refers to finding what proportions of materials to send from a destination to different processing locations (downstream processes) while respecting the mass flow conservation constraints. In addition, various production limit constraints related to the quality and quantity of materials mined, handled, processed, and sold also need to

154

be respected in the short-term production schedule. Like any complex network, uncertainty is inherent in a mining complex. This uncertainty stems from the quality, quantity, and spatial location of raw material within the mineral deposits – referred to as supply uncertainty – and from the production capabilities of different machinery, destinations, and processing streams – referred to as equipment performance uncertainty. A set of equally probable scenarios/simulations are generated using stochastic simulation methods to quantify supply uncertainty and variability (Goovaerts, 1997; Godoy, 2002; Boucher and Dimitrakopoulos, 2009; Remy et al., 2009), and equipment performance uncertainty (Quigley and Dimitrakopoulos, 2019).

A mining complex collects new information during its day-to-day operations with conventional and new digital technologies, specifically advanced sensors and monitoring devices. The new information acquired can pertain to the flow of materials (Rosa et al., 2007), equipment location (Chaowasakoo et al., 2014), equipment production capabilities (Koellner et al., 2004; Kargupta et al., 2010), and the quality and quantity of the material extracted, handled, and processed (De Jong, 2004; Goetz et al., 2009; Death et al., 2009; Iyakwari et al., 2016; Dalm et al., 2017, 2018; Kern et al., 2019). This new information provides an opportunity to better understand the state of a mining complex and to respond accordingly by adapting the short-term production schedule quickly (in real-time) to better meet the long-term targets. However, the core short-term scheduling decisions, such as the extraction sequence and destination policies, comprise a difficult combinatorial optimization problem and are computationally expensive to reoptimize with existing techniques (Lamghari, 2017). In addition, the incoming new information is partial and noisy, and is, therefore, uncertain. Thus, the new information cannot be used directly in an

155

optimization model to make short-term production scheduling decisions. The uncertain incoming new information needs to be assimilated to update the material supply and equipment performance uncertainties. Updating the supply uncertainty is more challenging compared to the equipment performance uncertainty due to the presence of multivariate spatial correlation. Ensemble Kalman filter is a well-known, two-step assimilation process that has been used to assimilate uncertain new information in petroleum reservoirs for decades (Aanonsen et al., 2009; Oliver and Chen, 2011) and has, recently, been extended to mineral deposits (Benndorf, 2015; Wambeke et al., 2018).

AI agents, such as deep neural networks and convolution neural networks, are function approximators that are trained to make decisions by responding to the incoming new information/observations/states (Sutton and Barto, 2017) via reinforcement learning algorithms that use their own experiences generated by interacting with an environment (a model that mimics the intricacies of the industrial process under consideration). The updated supply and equipment performance uncertainties allow such an AI agent to better perceive the updated state of a mining complex to adapt the short-term production schedule in real-time. The adapted production schedule is then used to perform day-to-day operations. In parallel, a new AI agent is trained with the updated supply and equipment performance uncertainties to further learn from the incoming new information. Therefore, a continuous updating framework is necessary, which first updates the uncertainties with the incoming uncertain new information, and then learns and adapts the short-term production schedule of a mining complex with AI agents. Benndorf and Buxton (2016) and Hou et al. (2015) proposed a framework that updates the supply uncertainty with incoming

new information but relies on existing optimization techniques for adapting the production scheduling decisions. Paduraru and Dimitrakopoulos (2019) proposed a policy gradient reinforcement learning algorithm for deciding the short-term destination of materials in a single product mining complex. However, the method does not adapt the extraction sequence, destination policies and processing stream utilization decisions simultaneously and does not update the supply and equipment performance uncertainties.

The work presented herein proposes a novel self-play reinforcement learning algorithm that adapts all the short-term production scheduling decisions simultaneously in a mining complex. The proposed algorithm is inspired by the AlphaGo and AlphaGoZero algorithms (Silver et al., 2016, 2017). The proposed algorithm plays the game of short-term production scheduling by itself using a Monte Carlo tree search to train a deep neural network agent to learn how to adapt the short-term production schedule with incoming new information in an operating mining environment. Additionally, the work also proposes a Monte Carlo simulation algorithm to update the equipment performance uncertainty. In the following sections, the proposed method that learns and adapts short-term decisions in a mining complex is detailed first. Next, an application at a copper mining complex is presented to show the efficiency of the proposed algorithm in terms of learning and adapting the short-term production schedule to generate more metal and achieve improved compliance with production targets. Conclusions and directions for future research follow.

4.2 Method

This section describes the algorithm for learning short-term production scheduling and then adapting the short-term production schedules with incoming new information in an operating mining environment. Section 4.2.1 details the process of transforming raw material into sellable products in a mining complex. The algorithm for learning and adapting the short-term production schedule is detailed next. A list of the notations used in this section is available in Appendix 4.1. The pseudo-code for the different parts of the proposed algorithm is provided in Appendix 4.3-Appendix 4.6.

4.2.1 Modelling a mining complex

Raw material in a mining complex can be supplied from various sources, such as multiple mines, $m \in M$. A mine is developed within the mineral deposit to extract materials. The materials in the mines consist of revenue-generating properties, $\mathbb{P}_{\mathbb{R}}$, deleterious properties, $\mathbb{P}_{\mathbb{D}}$, and rock mass, $\mathbb{P}_{\mathbb{M}}$. The material in the mines is discretized into a set of three-dimensional volumes called mining blocks, $\mathbb{z} \in \mathbb{Z}_m(x)$, where x denotes the spatial location of the block within the mine m. The quality of material in the mines is uncertain, therefore a set of initial, I, stochastic simulations, $\mathbb{s}_{I,a,m} \in \mathbb{S}_{I,a,m}$, of mining blocks, $\mathbb{Z}_{a,m}^{I,\mathbb{s}}(x)$, about multiple spatial correlated properties $a \in \mathbb{P}_{\mathbb{R}} \cup \mathbb{P}_{\mathbb{D}}$ is generated based on existing drill hole information $dH_{a,m}^I$ to quantify the supply uncertainty for mine $m \in M$. Each mine has a set of associated shovels, $\mathcal{S}_m$. The material extracted with the shovels is loaded into trucks, $\mathcal{T}_m$, available at each mine. The trucks haul the materials to their destinations, $\mathcal{D}$. The materials from the destinations are then transported via conveyor belts to processing streams, $\mathcal{P}$, to

generate products, $\mathbb{P}_\mathbb{R}$, which are then transported and sold to customers/markets. The processing streams have restrictions about the quantity of deleterious properties, $\mathbb{P}_\mathbb{D}$ in the final product. In addition, the trucks, shovels, destinations, and processing streams, $\mathbb{E} = \{\mathcal{T}, \mathcal{S}, \mathcal{D}, \mathcal{P}\}$, have restrictions on their production capacity, $\mathbb{P}_{\mathbb{P},e}(T), \forall e \in \mathbb{E}, T \in [1, N_{week}])$. N_{week} denotes the total number of weeks. A set of initial stochastic simulations, $s'_{I,e} \in \mathbb{S}'_{I,e}(T), \forall e \in \mathbb{E}, T \in [1, N_{week}]$, of production capabilities is generated based on historical production information eP_e^I to quantify the performance uncertainty of components $e \in \mathbb{E}$. The initial supply, $\mathbb{S}_{I,a,m}$, and equipment performance uncertainties, $\mathbb{S}'_{I,e}(T)$, are used in a self-play reinforcement learning algorithm to train a deep neural network (DNN) agent (see Sect. 4.2.2). The trained DNN agent is then used to make the short-term production scheduling decisions in real-time with the incoming new information (see Sect. 4.2.3).

The first short-term production scheduling decision in a mining complex is to decide when to extract a mining block. However, the multiple shovels located within the mine operate simultaneously; therefore, this decision variable is modified in such a way to take this into account. Let the blocks that can be extracted by a shovel be denoted by a set $\mathbb{Z}_m(s_i), \forall s_i \in \mathcal{S}_m, m \in M$. Let B represent a set whose elements are computed by finding all possible combinations that have exactly one element from each set $\mathbb{Z}_m(s_i), \forall s_i \in \mathcal{S}_m, m \in M$. The second short-term production scheduling decision in a mining complex is to decide the destination of the extracted block. Let $\mathcal{D}(z), \forall z \in \mathbb{Z}_m(x), m \in M$ denote the set of permissible destinations where each mining block can be sent. Therefore, the decision variable/action is defined by $x_{b,d,t}, \forall b \in \mathcal{B}, d \in \{\mathcal{D}(i), \forall i \in b\}$ and represents whether (1) or not (0) a set of

159

mining blocks b is extracted, and each block within the set b is sent to a set of destinations d at time step t.

$$x_{b,d,t} \leq x_{k,d',t-1} \quad , \forall k \in b_V, d' \in \{D(k), \forall k \in b_V\}, b \in B, d \in \{D(i), \forall i \in b\} \ (4.1)$$

In order to extract a block safely, it is necessary to extract all its overlying blocks to ensure the stability of the mine wall, and at least one surrounding block to provide access to the equipment. The vertical predecessor $V(i)$ defines all the mining blocks that overlie a block i, and is calculated by finding blocks that are within a predefined vertical inclination (known as the slope angle), as shown by the dashed lines in Figure 4-1. Let $b_V = \{V(i), \forall i \in b\}$ be a set consisting of all the blocks that lie above the blocks in a set b. Therefore, an action $x_{b,d,t}$ is only eligible to be made if all the overlying blocks (represented by decisions $x_{k,d',t-1}, \forall k \in b_V$) are extracted first, as shown in Eq. 4.1.

Figure 4-1 Permissible block extraction representation in a mining operation

Horizontal and vertical successor mining blocks define mine equipment access constraints. Horizontal successors $H(i,r)$ and vertical successors $V(i,r)$ for a block

$i \in b$, define all the surrounding blocks within a given radius r of a block i in the horizontal (solid line in Figure 4-1) and vertical directions (dotted line in Figure 4-1), respectively.

$$x_{b,d,t} \leq \sum_{k \in b_H} x_{k,d',t-1}, \forall k \in b_H, d' \in \{D(k)\}, b \in B, d \in \{D(i), \forall i \in b\} \quad (4.2)$$

Let b_H be a set consisting of k sets, where each set k consists of one of the blocks that surround a block i in the set b within a given radius r in either the horizontal or vertical directions. Therefore, an action $x_{b,d,t}$ is permissible only if at least one of the horizontal or vertical successors blocks (represented by decisions $x_{k,d',t-1}, \forall k \in b_H$) is extracted first, as shown in Eq. 4.2. The blocks that satisfy Eq. 4.1 and 4.2 also need to satisfy material classification conditions related to their destination, because the processing streams that are fed by destinations are designed to process a specific type of material. For example, a sulphide ore processor, by the construction of its metal recovery process, cannot accept oxide type materials. Therefore, a material classification condition is defined that finds the permissible destinations for a mining block. For example, if a mining block i has more soluble copper, then the permissible destination for this block can either be a destination that feeds the oxide processing stream or a waste dump. The permissible destinations of a mining block under supply uncertainty are determined as the most probable destinations (ties are broken randomly). The materials at the destinations incur a cost denoted by $C_{a,d}, \forall a \in \mathbb{P}_\mathbb{M}, d \in \mathcal{D}$ and are then sent to the processing streams. The third short-term production scheduling decision is to determine how to utilize the processing streams

in a mining complex, represented by $y_{a,d,\wp,t,\mathbb{s}_I^{\mathbb{J}}}$. The decision variable $y_{a,d,\wp,t,\mathbb{s}_I^{\mathbb{J}}}$ denotes the amount of attribute $a \in \mathbb{P}_{\mathbb{R}} \cup \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}$ sent from a destination $d \in \mathcal{D}$ to a processing location $\wp \in \mathcal{P}$ at time step t under joint uncertainty scenario $\mathbb{s}_I^{\mathbb{J}} \in \mathbb{S}_I^{\mathbb{J}}$. Here, joint uncertainty scenarios refer to uncertainty in both the supply of materials, $\mathbb{S}_{I,a,m}$, and performance of equipment, $\mathbb{S}'_{I,e}$. The processing streams recover the metal with a factor of $r_{a,\wp}, \forall a \in \mathbb{P}_{\mathbb{R}}, \wp \in \mathcal{P}$ and incur a processing cost of $C_{a,\wp}, \forall a \in \mathbb{P}_{\mathbb{M}}, j \in \mathcal{P}$. The products are then transported and sold to the customers with a price of $P_{a,\wp}, \forall a \in \mathbb{P}_{\mathbb{R}}, \wp \in \mathcal{P}$.

4.2.2 A self-play reinforcement learning algorithm

The algorithm proposed to learn to adapt the short-term production schedule (extraction sequence, destination policies, and processing stream utilization decisions, simultaneously) of a mining complex consists of a deep neural network (DNN) agent f_θ with parameters θ and a Monte Carlo tree search (MCTS) within a self-play reinforcement learning architecture. This type of algorithm is called a model-based deep reinforcement learning (DRL) algorithm. The motivation behind using a model-based DRL algorithm is to allow the algorithm to plan better by looking ahead in the future by accessing the environment. The proposed algorithm starts at time $t = 1$ with the input state s_1 of a mining complex. The state s_t of a mining complex at any time step t is described by the performance of its different components and is herein defined by (i) the position of the different shovels (the mining blocks that the shovels are extracting) located in the multiple mines, (ii) the supply uncertainty $\mathbb{S}_{I,a,m}$ of blocks located within a neighbourhood of the different shovels, (iii) the equipment

162

performance uncertainty of different components $\mathbb{S}'_{l,e}$, and (iv) history about the quality and quantity of material at destinations and processing streams (see Sect. 4.2.2.1 on how the history is generated). At each state, a set of permissible actions $x_{b,d,t}$ is identified using the mine wall slope stability, equipment access, and material classification criteria defined in Sect. 4.2.1. The state s_t is then fed to a DNN agent which outputs both a vector of selection probabilities and a vector of scalar evaluations $(p_t, v_t) = f_\theta(s_t)$ for all the permissible actions $x_{b,d,t}$ with the given input state s_t, as shown in Figure 4-2(b).

Figure 4-2 Self-play reinforcement learning algorithm for short-term production scheduling in mining complexes

For each time step $t < T$ an MCTS search α_θ is executed using tree policy P_π (see Sect. 4.2.2.1), guided by the DNN agent f_θ until the end time step T. The MCTS outputs probabilities π_t of selecting each action and scalar evaluations z_t for each action from the set of permissible actions at time step t as shown in Figure 4-2(a). An action $x_{b,d,t}$ at time step t is selected proportional to $\pi_t(x_{b,d,t}|s_t)$ (selection probabilities). These tree search probabilities π_t select much stronger actions $x_{b,d,t}$ than the raw probabilities p_t of the DNN agent $f_\theta(s_t)$. At this point, the DNN agent's

163

parameters θ are updated to make its selection probabilities and scalar evaluations $(p_t, v_t) = f_\theta(s_t)$ match the improved search probabilities π_t and evaluations z_t more closely (see Figure 4-2(b)). These new parameters make the MCTS search even stronger in the next round of self-play. The selected action $x_{b,d,t}$ is then used to update the input state as s_{t+1} and the process is repeated until $t = T, \forall\, T \in [1, N_{week}]$.

4.2.2.1 *Monte Carlo tree search*

The MCTS uses the DNN agent f_θ to guide its search. The search involves generating a rollout simulation of a feasible (that satisfies Eqs. 4.1 and 4.2) short-term production schedule (extraction sequence, destination policies, and processing stream utilization).

Figure 4-3 Monte Carlo tree search phases for short-term production scheduling in a mining complex. (a) selection; (b) expansion; (c) evaluation and simulation; (d) backup; and (e) playing an action (executed by repeating a-d)

For simplicity and without loss of generality, consider next a case where two shovels are located at the extreme ends of the mine (represented by a 2-dimensional grid), as shown in Figure 4-3(a). The colour of the grid cell indicates whether the

block is extracted (coloured) or not (white) and which destination it is sent to after it is extracted (red –crusher, green – leach, yellow – waste, and white – not extracted). Each node in the search tree represents a state s_t, and contains edges which represent a state-action pair $s_t, x_{b,d,t}$ where $x_{b,d,t}$ is a permissible child/action of the node s_t. Each edge in the search tree stores a set of statistics as prior probability $P(s_t, x_{b,d,t})$, visit count $N(s_t, x_{b,d,t})$, total action value $W(s_t, x_{b,d,t})$ and mean action-value $Q(s_t, x_{b,d,t})$. Multiple roll-out simulations are executed by iterating over four phases (a-d in Figure 4-3) and then selecting an action proportional to the search probabilities $\pi_t(x_{b,d,t}|s_t) = N(s_t, x_{b,d,t})/N(s_{t-1}, x_{b,d,t-1})$ (Figure 4-3(e)). The details of the MCTS in terms of selection, expansion, evaluation, simulation, and play phase are described next.

Selection

The selection phase (Figure 4-3(a)) is the first step in the tree search, which begins at the root node, i.e. in the input state s_1 of the MCTS, and finishes when the leaf node s_L at time step L is encountered. At each of these time steps $t < L$, an action $(x_{b,d,t})$ is selected that maximizes the upper confidence bound, $Q(s_t, x_{b,d,t}) + U(s_t, x_{b,d,t})$, calculated using the statistics stored in the tree search, as shown in Eq. 4.3, a variant of the PUCT algorithm (Silver et al., 2017).

$$x_{b,d,t} = \underset{x}{\mathrm{argmax}}\left(Q(s_t, x_{b,d,t}) + U(s_t, x_{b,d,t})\right) \qquad (4.3)$$

$$U(s_t, x_{b,d,t}) = c_{puct}P(s_t, x_{b,d,t})\frac{\sqrt{N(s_{t-1}, x_{b,d,t-1})}}{1 + N(s_t, x_{b,d,t})} \qquad (4.4)$$

Here, $U(s_t, x_{b,d,t})$ is calculated using Eq. 4.4 that uses the prior probability $P(s_t, x_{b,d,t})$ generated by the DNN agent and the visit count statistics $N(s_{t-1}, x_{b,d,t-1})$ stored in the MCTS for state-action pairs $s_t, x_{b,d,t}$. c_{puct}, is a constant that determines the level of exploration in the search, such that the search prefers actions with a high prior probability and a low visit count, but asymptotically prefers actions with high mean action-value.

Expansion

The leaf node s_L encountered during the selection phase of MCTS is then expanded (Figure 4-3(b)) to find the next permissible actions that satisfy Eqs. 4.1 and 4.2 which are then added to the search tree as children of node L. The different statistics of the edges $s_L, x_{b,d,L}$ connecting node s_L to its children nodes $x_{b,d,L}$ are initialized to zero, i.e. $N(s_L, x_{b,d,L}) = 0, Q(s_L, x_{b,d,L}) = 0, P(s_L, x_{b,d,L}) = 0, P(s_L, x_{b,d,L}) = 0$.

Evaluation and rollout simulation

The leaf node s_L state-action pair $s_L, x_{b,d,L}$ is then evaluated in two ways (Figure 4-3(c)), first by the DNN agent, and second by performing a rollout simulation until the end of time T using a tree policy P_π in MCTS. In comparison to playing games, the evaluation of short-term production scheduling in a mining complex (see Sect. 4.2.1) is more intricate. For instance, in the game of Go, the outcome is either a win or a loss but, in a mining complex, the outcome is an expected monetary gain from selling the products minus any losses and costs incurred to produce the products. In addition, the evaluation from the DNN agent requires some inputs about the history,

i.e. the quantity and quality of materials at the destination and the processing streams (see Sect. 4.2.2). The history of the quantity of attribute $a \in \mathbb{P}_{\mathbb{R}} \cup \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}$ at time step L at different destinations d in a state s_L is calculated by observing all the actions taken to reach node $x_{b,d,L}$.

$$v_{a,d,L,s_I} = \sum_{t<L} \sum_{b' \in b} \mathbb{Z}_{a,m}^{I,s}(b') \cdot \mathbb{Z}_{\mathbb{P}_{\mathbb{M}},m}^{I,s}(b') \cdot x_{b,d,t} \;, \forall d \in \mathcal{D}, s_I \in \mathbb{S}_{I,a,m}, a$$

$$\in \mathbb{P}_{\mathbb{R}} \cup \mathbb{P}_{\mathbb{D}} \tag{4.5}$$

$$v_{\mathbb{P}_{\mathbb{M}},d,L,s_I} = \sum_{t<L} \sum_{b' \in b} \mathbb{Z}_{a,m}^{I,s}(b') \cdot x_{b,d,t} \;, \forall d \in \mathcal{D}, s_I \in \mathbb{S}_{I,a,m} \tag{4.6}$$

Equation 4.5 calculates the quantity v_{a,d,L,s_I} of attribute $a \in \mathbb{P}_{\mathbb{R}} \cup \mathbb{P}_{\mathbb{D}}$ at destination $d \in \mathcal{D}$ at time step L under supply uncertainty $s_{I,a,m} \in \mathbb{S}_{I,a,m}$. Here, $\mathbb{Z}_{a,m}^{I,s}(b')$ represents the quantity of attribute a of a mining block b' at mine m in the initial I supply uncertainty scenario s. Equation 4.6 calculates the mass of materials at different destinations at time step L. The history of the quantity of attribute $a \in \mathbb{P}_{\mathbb{R}} \cup \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}$ at processing stream $\wp \in \mathcal{P}$ at time step L is calculated by optimizing the decision variable $y_{a,d,\wp,t,s_I^{\mathbb{J}}}$. For this, a stochastic mathematical programming model is solved using the simplex method that maximizes the objective function $f_L\left(s_I^{\mathbb{J}}\right)$, as shown in Eq. 4.7.

$$f_L\left(s_I^{\mathbb{J}}\right) = \sum_{\wp \in \mathcal{P}} \sum_{a \in \mathbb{P}_{\mathbb{R}}} P_{a,\wp} \cdot v_{a,\wp,L,s_I^{\mathbb{J}}} \cdot r_{a,i} - \sum_{\wp \in \mathcal{P}} \sum_{a \in \mathbb{P}_{\mathbb{M}}} C_{a,\wp} \cdot v_{a,\wp,L,s_I^{\mathbb{J}}}$$

$$- \sum_{\wp \in \mathcal{P}} \sum_{a \in \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}} \left(c_{a,\wp}^{+} \cdot d_{a,\wp,L,s_I^{\mathbb{J}}}^{+} + c_{a,\wp}^{-} \cdot d_{a,\wp,L,s_I^{\mathbb{J}}}^{-} \right) \;, \forall s_I^{\mathbb{J}} \in \mathbb{S}_I^{\mathbb{J}} \tag{4.7}$$

subjected to:

$$v_{a,\wp,L,\mathbb{s}_I^{\mathbb{J}}} - d^+_{a,\wp,L,\mathbb{s}_I^{\mathbb{J}}} \leq U_{a,\wp,T} \quad , \forall\, a \in \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}, \wp \in \mathcal{P}, \mathbb{s}_I^{\mathbb{J}} \in \mathbb{S}_I^{\mathbb{J}} \tag{4.8}$$

$$v_{a,\wp,L,\mathbb{s}_I^{\mathbb{J}}} + d^-_{a,\wp,L,\mathbb{s}_I^{\mathbb{J}}} \geq L_{a,\wp,T} \quad , \forall\, a \in \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}, \wp \in \mathcal{P}, \mathbb{s}_I^{\mathbb{J}} \in \mathbb{S}_I^{\mathbb{J}} \tag{4.9}$$

$$\sum_{\wp \in \mathcal{P}} y_{a,d,\wp,L,\mathbb{s}_I^{\mathbb{J}}} = 1 \quad , \forall d \in \mathcal{D}, \mathbb{s}_I^{\mathbb{J}} \in \mathbb{S}_I^{\mathbb{J}} \tag{4.10}$$

where,

$$v_{a,\wp,L,\mathbb{s}_I^{\mathbb{J}}} = \sum_{d \in \mathcal{D}} y_{a,d,\wp,L,\mathbb{s}_I^{\mathbb{J}}} \cdot v_{a,d,L,\mathbb{s}_I} \quad , \forall\, a \in \mathbb{P}_{\mathbb{R}} \cup \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}, \wp \in \mathcal{P}, \mathbb{s}_I^{\mathbb{J}} \in \mathbb{S}_I^{\mathbb{J}} \tag{4.11}$$

The objective function (Eq. 4.7) computes the profit generated by processing the materials, minus the cost of processing the material and penalties due to deviations from the upper $U_{a,\wp,T}$ and lower $L_{a,\wp,T}$ production capacities. Here, $c^+_{a,\wp}$ and $c^-_{a,\wp}$ is the penalty cost for deviating from the upper and lower production capacities, respectively. $v_{a,\wp,L,\mathbb{s}_I^{\mathbb{J}}}$ is the quantity of attribute a at processing location $\wp$ under joint uncertainty scenario $\mathbb{s}_I^{\mathbb{J}}$ at time step L. Equations 4.8 and 4.9 constrain the quantity of attribute $a \in \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}$ at processing location $\wp \in \mathcal{P}$ under joint uncertainty scenario $\mathbb{s}_I^{\mathbb{J}} \in \mathbb{S}_I^{\mathbb{J}}$ within the upper and lower production capacities while allowing for deviations $d^+_{a,\wp,L,\mathbb{s}_I^{\mathbb{J}}}$ and $d^-_{a,\wp,L,\mathbb{s}_I^{\mathbb{J}}}$ from such capacities, respectively. Equation 4.10 ensures that mass flow balancing is conserved while solving the stochastic optimization model. The history of materials at different destinations and processing streams, along with other information mentioned in Sect. 4.2.2, is then fed to the DNN

agent to generate both prior probabilities p_L and scalar evaluations v_L for nodes $x_{b,d,L}$. The second evaluation is the rollout simulation until the time step T using a tree policy P_π. The rollout policy P_π consists of, at each time step $t \in [L, T]$, (i) finding permissible blocks to extract using Eqs. 4.1 and 4.2, (ii) finding permissible block destinations using cut-off grade policies (Lane, 1988; Rendu, 2014), (iii) combining the two to form permissible actions $x_{b,d,t}$ for state s_t, (iv) randomly selecting an action for state s_t among the set of permissible actions, and then (v) optimizing the processing stream utilization decisions with the stochastic mathematical programming model (Eqs. 4.7-4.11). Equations 4.5 and 4.6 are then used to update the history of materials at destinations at time step T with the actions $x_{b,d,t}, \forall t \in [L, T]$ selected during the Monte Carlo tree search.

$$v_{a,d,T,\mathbb{s}_I} - d^+_{a,d,T,\mathbb{s}_I^\mathbb{J}} \leq U_{a,d,T} \quad , \forall \, a \in \mathbb{P}_\mathbb{M}, d \in \mathcal{D}, \mathbb{s}_I^\mathbb{J} \in \mathbb{S}_I^\mathbb{J} \qquad (4.12)$$

$$v_{a,d,T,\mathbb{s}_I} + d^-_{a,d,T,\mathbb{s}_I^\mathbb{J}} \geq L_{a,d,T} \quad , \forall \, a \in \mathbb{P}_\mathbb{M}, d \in \mathcal{D}, \mathbb{s}_I^\mathbb{J} \in \mathbb{S}_I^\mathbb{J} \qquad (4.13)$$

Equations 4.12 and 4.13 are used to calculate the deviations $d^+_{a,d,T,\mathbb{s}_I^\mathbb{J}}, d^-_{a,d,T,\mathbb{s}_I^\mathbb{J}}$ from upper and lower production capacities, respectively, at the destinations at time step T. Here, $U_{a,d,T}$ is the $\max\left(\mathbb{S}'_{I,d}(T)\right)$ and $L_{a,d,T}$ is the $\min\left(\mathbb{S}'_{I,d}(T)\right), \forall a \in \mathbb{P}_\mathbb{M}, d \in \mathcal{D}$.

$$r_T = \frac{1}{|\mathbb{S}_I^\mathbb{J}(T)|} \underbrace{\sum_{s \in \mathbb{S}_I^\mathbb{J}(T)} \sum_{\wp \in \mathcal{P}} \sum_{a \in \mathbb{P}_\mathbb{R}} P_{a,\wp} \cdot v_{a,\wp,T,s} \cdot r_{a,\wp}}_{\text{Part I}}$$

$$-\frac{1}{|\mathbb{S}_I^{\mathbb{J}}(T)|} \sum_{s \in \mathbb{S}_I^{\mathbb{J}}(T)} \sum_{i \in \mathcal{P} \cup \mathcal{D} \cup M} \sum_{a \in \mathbb{P}_{\mathbb{M}}} C_{a,i} \cdot v_{a,i,T,s}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{Part II}}$$

$$-\frac{1}{|\mathbb{S}_I^{\mathbb{J}}(T)|} \sum_{s \in \mathbb{S}_I^{\mathbb{J}}(T)} \sum_{i \in \mathcal{P} \cup \mathcal{D}} \sum_{a \in \mathbb{P}_{\mathbb{D}} \cup \mathbb{P}_{\mathbb{M}}} \left(c_{a,i}^+ \cdot d_{a,i,T,s}^+ + c_{a,i}^- \cdot d_{a,i,T,s}^- \right) \qquad (4.14)$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{Part III}}$$

The stochastic mathematical programming model (Eqs. 4.7-4.11) is used to compute $y_{a,d,\wp,T,\mathbb{S}_I^{\mathbb{J}}}$ and deviations $d_{a,\wp,L,\mathbb{S}_I^{\mathbb{J}}}^+ , d_{a,\wp,L,\mathbb{S}_I^{\mathbb{J}}}^-$ at time step T. Finally, the expected future reward r_T under joint uncertainty $\mathbb{S}_I^{\mathbb{J}}(T)$ for the rollout simulation is computed using Eq. 4.14. Part I of Eq. 4.14 represents the profit from selling all the products, Part II includes all the costs incurred to generate the products, such as mining, crushing, stockpiling, and processing costs, and Part III represents the penalties for deviating from the different production limits.

Backup

The last phase of the tree search is the backup phase (Figure 4-3(d)) where first the visit count of all nodes s_t, visited at each time step $t \leq L$ until the leaf node was reached in the selection phase, is increased by 1 using Eq. 4.15 shown below.

$$N(s_t, x_{b,d,t}) = N(s_t, x_{b,d,t}) + 1 \qquad , \forall t \leq L \qquad (4.15)$$

$$W(s_t, x_{b,d,t}) = W(s_t, x_{b,d,t}) + r_t \qquad , \forall t \leq L \qquad (4.16)$$

$$Q(s_t, x_{b,d,t}) = \frac{W(s_t, x_{b,d,t})}{N(s_t, x_{b,d,t})} (1 - \gamma) + \gamma \cdot v_t \qquad , \forall t \leq L \qquad (4.17)$$

The total action value of each node s_t visited at each time step $t \leq L$ is updated using Eq. 4.16. The mean action value for each node s_t visited at each time step $t \leq L$ is then updated by mixing the scalar evaluations from the DNN agent with a factor γ, and the mean action value stored in the search tree with a factor $1 - \gamma$, as shown in Eq. 4.17.

Play

The MCTS search is repeated for N_{MCTS}^{Train} times and then finally an action $x_{b,d,t}$ at time step t is selected (Figure 4-3(e)), proportional to the visit counts of all actions $x_{b,d,t}$, i.e. $\pi_t(x_{b,d,t}|s_t) \propto N(s_t, x_{b,d,t})/N(s_{t-1}, x_{b,d,t-1})$. The selected action is used in Eqs. 4.5-4.11 to update the state to s_{t+1}. The selected action becomes the new root for the next round of self-play, and the process continues until terminal time step T is reached.

4.2.2.2 *DNN agent training*

The DNN agent f_{θ_t} is initialized at time $t = 1$ with random weights θ_1. At each subsequent time step $t \leq T$, an MCTS search $\alpha_{\theta_{t-1}}(s_t)$ is executed using the DNN agent from the previous step $f_{\theta_{t-1}}$ and tree policy P_π to return both the search probabilities π_t and search scalar evaluations z_t.

$$l = ||z - v||^2 - \pi^{\mathrm{T}} \log p + c||\theta||^2 \tag{4.18}$$

The data for each time-step t is stored as (s_t, π_t, z_t) and used to locally train the DNN agent to generate new parameters f_{θ_t}, for N_L iterations. More specifically, the DNN parameter f_θ is adjusted by stochastic gradient descent on a loss function (l) that minimizes the cross-entropy loss between action probabilities p and search probabilities π, and the mean-squared error between scalar predicted v and search evaluations z as shown in Eq. 4.18. An L2 regularization with penalty cost c is added in the loss function to avoid overfitting. At time step $t = T$, all the stored training data prior to time step t is then used to train the DNN agent globally for N_G iterations to avoid overfitting to any specific data instances generated from any specific time step. The new DNN agent is then used for further rounds of self-play and training.

4.2.3 Responding to incoming new information

The self-play reinforcement learning algorithm (Sect. 4.2.2) generates a DNN agent that can adapt the short-term production schedule of a mining complex with incoming new information. The new information collected during mining operations is first used to generate updated supply $\mathbb{S}_{U,a,m}$ and equipment performance uncertainties $\mathbb{S}'_{U,e}$. The updated uncertainties are then fed to the trained DNN agent to adapt the short-term production schedule, which is then used to generate the updated production forecasts. In parallel, a new DNN agent is trained with the updated uncertainties to further adapt the agent parameters. The algorithm is general and can be applied to different mining complexes.

4.3 Application at a copper mining complex

The proposed self-play reinforcement learning algorithm is applied at a copper mining complex. In the case study, the incoming new information about the supply of materials is the blasthole data measured during drilling at the mine, and the productivity data about the shovels, trucks, and crushers collected during the mine's operation. However, the algorithm is flexible enough to include different types of incoming new information related to the supply of materials and the performance of the different components of a mining complex.

4.3.1 Overview of the mining complex

The copper mining complex consists of two mines, mine A and mine B. The materials extracted with the multiple shovels at the two mines are transported via trucks to five different crushers, a waste dump, and a sulphide leach pad, represented by $\mathcal{D}$ as shown in Figure 4-4. The materials from the crushers are then transported to three different processing mills and an oxide leach pad via conveyor belts. The processing mills produce copper (Cu) concentrate as a primary product and gold (Au), silver (Ag), and molybdenum (Mo) concentrates as secondary/by-products. The leach pads supply materials to a copper cathode plant that generates copper plate products. The processing mills and the cathode plant are represented by $\mathcal{P}$. The products (concentrates and plates) are then transported and sold to different customers.

173

Figure 4-4 The Copper mining complex

Table 4-1 Material classification criteria and cut-off grade policies for copper mining complex

Materials classification	Materials classification criteria	Permissible destinations	Cut-off grade destination policies	Cut-off grade destination
High-grade Sulphide	$CuS/CuT \leq 0.2$	Processing mill, sulphide leach pad, and waste dump	$CuT \geq 0.6$	Processing mill
			$0.3 \leq CuT < 0.6$	Sulphide leach pad
			$CuT < 0.3$	Waste dump
Low-grade Sulphide	$0.2 < CuS/CuT \leq 0.5$	Processing mill, sulphide leach pad, and waste dump	$CuT > 0.3$	Sulphide leach pad
			$CuT \leq 0.3$	Waste dump
Oxide	$CuS/CuT \geq 0.5$	Oxide leach pad and waste dump	$CuS \geq 0.2$	Oxide leach pad
			$CuS < 0.2$	Waste dump

The proposed algorithm is used to adapt the weekly ($N_{week} = 13$) short-term production schedule of the copper mining complex within a given quarterly

174

production schedule which, for this case study, consists of 3,600 and 1,200 mining blocks from each of the two mines, respectively. The mining blocks have properties such as copper soluble (CuS), copper total (CuT), gold (Au), silver (Ag), and Molybdenum (Mo), which generate revenue, arsenic (As) which is deleterious, and the mass of the block. All the components of the mining complex have limits on their production capacity. Additionally, the processing mill has a limit on the amount of arsenic in the product. Table 4-1 shows the material classification criteria, permissible destinations, and cut-off grade policies used at the copper mining complex. The production limits of the different components of the mining complex are listed in Appendix 4.2. The economic and operational parameters used at the copper mining complex are also detailed in Appendix 4.2. Additional results of the case study are presented in Appendix 4.8.

4.3.2 Parameters

The proposed algorithm was run on an Intel® i7-8700 machine with an 8-core processor and an NVIDIA GeForce GTX 1050 GPU. The algorithm uses the Tensorflow Adam optimizer with default settings (Kingma and Ba, 2015) to train the DNN agent for approximately 2 days. The neighbourhood of blocks used as an input was set to 81 blocks for each shovel. The inputs to the DNN agent are normalized between 0 and 1. The scalar evaluations of the MCTS are also normalized between 0 and 1 for them to be of the same magnitude as the search probabilities. The number of next permissible actions is fixed to be 256. The missing actions are filled in by duplicating the existing actions until 256 is reached. If there are more than 256 permissible actions, then only the first 256 actions are considered. The training

175

process in this case study was executed with 2 local and 20 global iterations. Over the course of training, 1.9 million different weekly production schedules were generated to train the DNN agent. 500 (N_{MCTS}^{Train}) rollout simulations were performed for selecting an action in the MCTS, which corresponds to approximately 30s of think time. The mixing parameter, L2 regularization cost, and MCTS exploration parameter are set to 0.25, 0.01, and 2, respectively. Most of the parameters in the case study are set by a trial and error mechanism to generate stable solutions. The most sensitive parameters in the proposed algorithm are the mixing and MCTS exploration parameters. The rest of the parameters such as the number of rollout simulations, neighbourhood blocks, and number of permissible actions should be set based on the intricacy of the mining complex. Ten stochastic simulations for each of the two mines and all their equipment were used to train the DNN agent. The stochastic simulations of the mines were combined with stochastic simulations of the equipment to generate joint uncertainty scenarios. For example, 10 stochastic simulations of the mines and 10 stochastic simulations of the equipment result in 100 joint uncertainty scenarios. The architecture of the DNN used in the case study is presented in Appendix 4.7.

4.3.3   Results

Section 4.3.3.1 shows the result of updating the supply and equipment performance uncertainties with incoming new information, and Sect. 4.3.3.2 details the result of adapting the 13-week short-term production schedule. The results in this section are scaled for confidentiality reasons, with the forecast of an initial production schedule over the initial uncertainties being 100%.

4.3.3.1 *Updated supply and equipment performance uncertainties*

The incoming new information about the properties of materials $l_a^{NI}, \forall a \in \mathbb{P}_{\mathbb{R}} \cup \mathbb{P}_{\mathbb{D}}$ is used to generate the updated supply uncertainty $\mathbb{S}_{U,a,m}$ with an extended ensemble Kalman filter (EnKF) method that accounts for the multivariate spatial correlation of different properties. The details of the method can be found in Chapter 3. Supply uncertainty about six correlated properties, namely CuS, CuT, As, Au, Ag, and Mo, is updated with the new blasthole information (3,437 blasthole data) collected during the mine's operation. Figure 4-5 shows one of the initial and updated stochastic simulations of the CuT and As properties at bench 1 of mine A. The extended EnKF method updates the local characteristics of materials at mine A based on the new blasthole information. The concentration and spatial distribution of As have changed significantly in the updated simulations, as seen in Figure 4-5.

Figure 4-5 (a) Initial simulation updated with (b) incoming blasthole data to generate (c) updated simulation for bench 1 at mine A

177

The incoming new information about the productivity of different components $l_e^{NI}, \forall e \in \mathbb{E}$ is used to update the initial equipment performance uncertainty $\mathbb{s}'_{I,e} \in \mathbb{S}'_{I,e}(T), \forall e \in \mathbb{E}$ by first computing the empirical cumulative distribution function (ECDF) with both the historical data eP_e^I and incoming new information l_e^{NI}. The ECDF is then sampled to generate the updated, U, equipment performance uncertainty $\mathbb{s}'_{U,e} \in \mathbb{S}'_{U\,e}(T)$ of the different components $e \in \mathbb{E}$. Figure 4-6 shows the initial and updated simulations of the production capabilities of crusher 5. The production capabilities of crusher 5 in the initial and updated simulations are different but still respect the data. Updating the supply and equipment performance uncertainties with incoming new information takes about five minutes in this case study.

Figure 4-6 (a) Initial and (b) updated simulations about crusher 5 production capacity

The updated supply $\mathbb{S}_{U,a,m}$, and equipment performance uncertainties $\mathbb{S}'_{U,e}(T)$ are fed to the DNN agent that responds to the updated uncertainties by adapting the short-term production schedule (see Sect. 4.2.3). The results of the adapted short-term production schedule are discussed next.

4.3.3.2 *Adapted short-term production schedule*

The results in this section are reported using the 10ᵗʰ, 50ᵗʰ, and 90ᵗʰ percentile risk profiles (P10, P50, and P90, respectively) of the different performance indicators over a set of 100 joint uncertainty scenarios (separate from the ones used to train the DNN agent). Three different types of results are reported, (a) green lines: performance of an existing short-term production schedule under initial supply and equipment uncertainties, (b) black lines: risk of the initial short-term production schedule under the updated uncertainties generated in Sect. 4.3.3.1 (the block extraction sequence and destination decisions are fixed from the initial production schedule, and the processing stream utilization decisions are reoptimized using Eqs. 4.7-4.11), and (c) blue lines: performance of the adapted short-term production schedule generated by the DNN agent based on the updated uncertainties. The results indicated by (b) and (c) in Figure 4-9-Figure 4-13 show the value of updating the uncertainties and the added value of adapting the short-term production schedule (the ability of the DNN agent to respond to incoming new information), respectively. Although a mining operation will use this framework to adapt and learn continuously, for simplicity and fair comparison, no Monte Carlo tree searches or training are performed over the updated uncertainties. This algorithm takes only about two minutes to adapt the 13-week short-term production schedule in this case study.

Extraction sequence and destination of materials

Figure 4-7 and Figure 4-8 show the block extraction sequence and destination decisions, respectively, for bench 1 at mines A and B for the (a) initial and (b) adapted short-term production schedules. The initial and adapted short-term production

179

schedules respect the permissible block extraction and destination constraints. The adapted block extraction and destination decisions are very different from the initial ones. The DNN agent is efficient at responding to the updated uncertainties by adapting the extraction sequence to better blend the material and respect the production limits of the different components. Additionally, the destination of the blocks is adapted by the DNN agent to better utilize the processing capabilities and produce a higher quantity of primary products. The reasons for the major differences in the initial and adapted short-term production schedules are due to:

a. The ability of the proposed algorithm to capitalize on the synergies between the different components of the mining complex to simultaneously adapt all relevant short-term production scheduling decisions.

b. The ability of the proposed algorithm to account for multiple sources of uncertainty related to the supply of material and production capabilities of different components of the mining complex.

c. The ability of the continuous updating framework to allow the DNN agent to better observe the updated state of the mining complex, which includes uncertainty about revenue-generating and deleterious properties of materials and the production capabilities of its components.

Figure 4-7 Block extraction sequence in (a) the initial short-term production schedule compared to (b) the adapted short-term production schedule for bench 1

Figure 4-8 Block destination decisions in (a) the initial short-term production schedule compared to (b) the adapted short-term production schedule for bench 1

Ore production forecasts

The forecasts for the initial and adapted short-term production schedules for the different production limits are shown in this section. Figure 4-9(a) shows the performance of the initial short-term production schedule for the production capacity limit of mill 1. The production schedule respects the capacity limits of mill 1 with some weeks of lower utilization. Figure 4-9 (b) shows the risk profile of the initial production schedule over the updated uncertainties and presents a lower utilization of

181

the capacity of mill 1. Figure 4-9 (c) shows the performance of the adapted short-term production schedule generated by the DNN agent. The capacity of mill 1 is better utilized in the adapted production schedule.

Figure 4-9 Mill 1 production limit forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-10 Mill 1 arsenic limit forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-10(a), (b), and (c) show the forecasts for the initial short-term production schedule, its risk profile over the updated uncertainties, and the adapted schedule for the arsenic quality limit for mill 1, respectively. The arsenic quality limit is well respected until week 8 in the initial schedule (Figure 4-10(a)); however, the risk of the initial schedule over the updated uncertainties (Figure 4-10 (b)) shows that this

limit will be violated in the early weeks (because of a high arsenic concentration in the materials in mine A as seen from Figure 4-10(c)). The DNN agent adapts the short-term schedule to blend materials from multiple mines, and the violations are minimal in the adapted schedule (Figure 4-10(c)).

Metal production forecasts

Figure 4-11(a), (b), and (c) show the forecasts for copper concentrate production for the initial short-term production schedule, the risk of the initial schedule over the updated uncertainties, and the adapted schedule, respectively. The initial schedule realized over the updated uncertainties (Figure 4-11(b)) shows an increase of 2 % in recovered copper concentrate production. However, the DNN agent adapts the short-term schedule to increase the copper concentrate by 14%. The adapted schedule better blends the materials to better utilize the processing mill capacities and results in a higher copper concentrate production.

(a)   (b)   (c)

Figure 4-11 Copper concentrate forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-12 Gold concentrate forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-12 shows the forecasts for gold concentrate production. The initial short-term production schedule realized over the updated uncertainties shows an 11% increase in recovered gold production (Figure 4-12(b)). The adapted schedule by the DNN agent increases gold production by 22% (Figure 4-12 (c)). The adapted schedule processes less material with the leach pads to better blend the materials to meet the arsenic quality limits of the processing mills (Figure 4-10(c)), to better utilize the processing mill capacities (Figure 4-9(c)), and to generate a higher quantity of primary copper concentrate and secondary gold, silver and molybdenum products (Figure 4-11(c) and Figure 4-12(c)).

Cash flow forecasts

Figure 4-13(a), (b), and (c) show the cumulative cash flow forecasts of the initial short-term production schedule, the initial schedule realized over the updated uncertainties, and the adapted schedule, respectively.

Figure 4-13 Cumulative cash flow forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

The initial short-term production schedule shows an increase of 5% in cash flow when realized over the updated uncertainties, which shows the value of updating the uncertainties with incoming new information. However, the DNN agent adapts the short-term production schedule to generate a 12% increase in cash flow, i.e. an additional 7% of added value in adapting the short-term schedule. The added cash flow value in the adapted schedule is generated by a better extraction sequence and destination decisions, and improved blending strategies to maximize the utilization of the processing stream capacities.

4.4   Conclusions

This paper proposes a new self-play reinforcement learning algorithm that combines a Monte Carlo tree search with a deep neural network agent to adapt the short-term production schedule of a mining complex with incoming new information. The deep neural network agent evaluates the short-term production scheduling decisions and, in parallel, uses the evaluations in a Monte Carlo tree search to gather self-play experiences by performing random rollouts. The gathered experiences are then used to train the deep neural network agent, which improves the strength of the

tree search and results in stronger self-plays to generate better experiences. First, the incoming new information is used in the extended EnKF algorithm proposed in Chapter 3 and the Monte Carlo simulation algorithm proposed in this work to update the supply and equipment performance uncertainties. The updated uncertainties are then fed to the self-play reinforcement learning algorithm proposed in this work, to adapt all the relevant short-term production scheduling decisions (sequence of extraction, the destination of materials, and utilization of processing stream) in a mining complex simultaneously. An application of the proposed algorithm at a copper mining complex shows its exceptional performance in adapting the 13-week short-term production schedule with incoming new information. The risk profiles of realizing the initial 13-week short-term production schedule of the copper mining complex over the updated supply and equipment performance uncertainties showed an increase of 5% in cumulative cash flow, and an increase of 2%, 11%, 23%, and 32% in copper concentrate, gold, silver, and molybdenum production, respectively. It also showed a large violation of the arsenic content limit and a lowered utilization of processing capacities. The proposed self-play reinforcement learning algorithm adapted the short-term production schedule to increase the cumulative cash flow by 12%, and the copper concentrate, gold, silver, and molybdenum production by 14%, 22%, 43%, and 61%, respectively. In addition, the adapted short-term production schedule makes better use of the processing mill capacities and shows a minimal violation of the arsenic quality limit of the processing mills by finding better blending strategies, which result in a 2% reduced copper cathode production. The process of adapting the 13-week short-term production schedule of the copper mining complex takes two minutes in this case study. In the future, the research will focus on

186

integrating more sources of incoming new information to update the supply uncertainty and modifying the self-play reinforcement learning algorithm to use convolution neural network agents without a tree policy, like the AlphaGoZero algorithm.

Appendix 4.1

This section presents the notations used in the proposed self-play reinforcement learning algorithm. Table 4-2 shows the sets and indices used in the proposed algorithm. Table 4-3 and Table 4-4 present the decision variables and constants used in the proposed algorithm, respectively.

Table 4-2 Sets and indices used in the proposed algorithm

Parameters	Definition
M	Set of mines, $m \in M$
I	Initial
U	Updated
N_{week}	Number of weeks in the quarterly short-term production schedule
$\mathbb{P}_{\mathbb{R}}$	Set of revenue-generating properties attribute
$\mathbb{P}_{\mathbb{D}}$	Set of deleterious properties attribute
$\mathbb{P}_{\mathbb{M}}$	Rock mass attribute
$\mathbb{Z}_m(x)$	Set of mining blocks in mine m located at x, $\mathbb{z} \in \mathbb{Z}_m(x)$
$dH_{a,m}^I$	Initial drill hole information for attribute a in mine m
$\mathbb{S}_{I,a,m}$	Set of initial stochastic simulations for attribute a in mine m, $\mathbb{s}_{I,a,m} \in \mathbb{S}_{I,a,m}$
$\mathbb{Z}_{a,m}^{I,\mathbb{S}}(x)$	Property of a set of blocks located at x in mine m for attribute a in initial stochastic simulation $\mathbb{s}_{I,a,m}$
$\mathbb{S}_I^{\mathbb{J}}$	Set of initial joint uncertainty scenarios, $\mathbb{s}_I^{\mathbb{J}} \in \mathbb{S}_I^{\mathbb{J}}$, $\mathbb{S}_I^{\mathbb{J}} = \mathbb{S}_{I,a,m} \cup \mathbb{S}_{I,e}'(T), \forall\, T \in [1, N_{week}]$
$\mathcal{S}_m$	Set of shovels located in mine m, $\mathscr{s}_i \in \mathcal{S}_m$
$\mathcal{T}_m$	Set of trucks used in mine m
$\mathcal{D}$	Set of destinations in the mining complex
$\mathcal{P}$	Set of processing streams in the mining complex
$\mathbb{S}_{I,e}'(T)$	Set of initial stochastic simulations for component e in period T, $\mathbb{s}_{I,e}' \in \mathbb{S}_{I,e}'(T), \forall\, T \in [1, N_{week}]$
eP_e^I	Initial production information (historical information) for component e
$\mathbb{P}_{\mathbb{P},e}(T)$	Production capacity for component e in period T, $\forall e \in \mathbb{E}$, $\mathbb{E} = \{\mathcal{T}, \mathcal{S}, \mathcal{D}, \mathcal{P}\}, \forall\, T \in [1, N_{week}]$

$\mathbb{Z}_m(\mathcal{s}_i)$	Set of permissible blocks that can be extracted from mine m with the shovel $\mathcal{s}_i$
B	Set consisting of sets where each set has one element from each set in $\mathbb{Z}_m(\mathcal{s}_i)$
$D(\mathbb{z})$	Set of permissible destinations for each mining block $\mathbb{z}$
$\mathbb{S}'_{U\,e}(T)$	Set of updated stochastic simulations for e in period T, $\mathbb{s}'_{U,e} \in \mathbb{S}'_{U\,e}(T), \forall\, T \in [1, N_{week}]$
l_e^{NI}	Real-time new information about the performance of component e
l_a^{NI}	Real-time new information about the supply of material attribute a
$\mathbb{S}_{U,a,m}$	Set of updated stochastic simulations for attribute a in mine m, $\mathbb{s}_{U,a,m} \in \mathbb{S}_{U,a,m}$
$V(i)$	Set of blocks that are the vertical predecessor of block i
b'_V	Set of all the blocks that overlie all the blocks in set $b, \forall b \in B$
$H(i,r)$	Set of blocks that are horizontal successor (surrounding blocks) of block i within a radius r
b'_H	Set consisting of k sets, with each set k consist of one block that surrounds a block i in set b within a radius $r, \forall b \in B$
$v_{a,d,t,\mathbb{s}_I}$	Quantity of the attribute a at destination d at time step t under stochastic simulation $\mathbb{s}_{I,a,m}, \forall\, t \in [1, T]$
$v_{a,\wp,t,\mathbb{s}_I^{\mathbb{J}}}$	Quantity of the attribute a at processing stream $\wp$ at time step t under joint uncertainty scenario $\mathbb{s}_I^{\mathbb{J}}, \forall\, t \in [1, T]$
$L_{a,j,T}$	Lower limit for attribute a at location j in period $T, \forall\, T \in [1, N_{week}]$
$U_{a,j,T}$	Upper limit for attribute a at location j in period $T, \forall\, T \in [1, N_{week}]$

Table 4-3 Variables used in the proposed algorithm

Variable	Definition
$x_{b,d,t} \in \{0,1\}$	Defines if a set of blocks b is extracted from a set of shovels and sent to a set of destinations d at time step $t, \forall\, t \in [1, T]$
$y_{a,d,\wp,t,\mathbb{s}_I^{\mathbb{J}}} \in [0,1]$	Amount of attribute a sent from destination d to processing stream $\wp$ at time step t under joint uncertainty scenario $\mathbb{s}_I^{\mathbb{J}},\,, \forall\, t \in [1, T]$
f_θ	Deep neural network (DNN) agent with parameters θ
s_t	State of the mining complex at time $t, \forall\, t \in [1, T]$

p_t, v_t	Vector of selection probabilities and scalar evaluation for all the permissible actions $x_{b,d,t}$ in the state s_t, $(p_t, v_t) = f_\theta(s_t)$, $\forall\, t \in [1, T]$
α_θ	Monte Carlo tree search (MCTS)
π_t	Selection probabilities at time step t output from Monte Carlo tree search α_θ, $\forall\, t \in [1, T]$
z_t	Scalar evaluation at time step t from Monte Carlo tree search α_θ, $\forall\, t \in [1, T]$
$P(s_t, x_{b,d,t})$	The prior probability for state-action pair $(s_t, x_{b,d,t})$
$W(s_t, x_{b,d,t})$	Total action value for state-action pair $(s_t, x_{b,d,t})$
$Q(s_t, x_{b,d,t})$	Mean action value for state-action pair $(s_t, x_{b,d,t})$
$N(s_t, x_{b,d,t})$	Visit count for state-action pair $(s_t, x_{b,d,t})$
$d^+_{a,j,t,\mathbb{s}^{\mathbb{J}}_I}$	Continuous variable for deviation from the upper limit $U_{a,j,T}$ at time step t for attribute a at location j under joint uncertainty scenario $\mathbb{s}^{\mathbb{J}}_I$, $\forall\, t \in [1, T]$
$d^-_{a,j,t,\mathbb{s}^{\mathbb{J}}_I}$	Continuous variable for deviation from the lower limit $L_{a,j,T}$ at time step t for attribute a at location j under joint uncertainty scenario $\mathbb{s}^{\mathbb{J}}_I$, $\forall\, t \in [1, T]$
r_t	Total cumulative future expected reward from time step t, $\forall\, t \in [1, T]$

Table 4-4 Constants used in the proposed algorithm

Constants	Definition
$C_{a,d}$	Cost incurred for material attribute a at destination d
$r_{a,\wp}$	Recovery factor for attribute a at processing location $\wp$
$C_{a,\wp}$	Cost of processing material attribute a at processing location $\wp$
$P_{a,\wp}$	Price of selling material attribute a at processing location $\wp$
P_π	Sub-optimal tree policy
c_{puct}	A factor that determines the level of exploration in the selection phase of MCTS
γ	Mixing parameter for neural network policy reward and MCTS reward
c	L2 regularization factor with neural network
N_L	Number of local epochs for training DNN agent
N_G	Number of global epochs for training DNN agent
N_{MCTS}^{Train}	Number of MCTS simulations in the training phase

N_{MCTS}^{Update}	Number of MCTS simulations in adapting phase
A	Vertical Slope angle
r	Horizontal radius

Appendix 4.2

This section presents the production limits and operational and economic parameters of the copper mining complex. The production limits of the different components of the mining complex are presented in Table 4-5. The operational and economic parameters used in the case study are outlined in Table 4-6. The production targets and economic parameters are scaled for confidentiality purposes.

Table 4-5 Production limits of the different components of the copper mining complex

Attribute	Value (Weekly)
Crusher 1, 2, 3, 4, and 5 production capacity limit (Tonnes)	Stochastic
Mill 1, 2, and, 3 capacity limit (Tonnes)	28, 33.5, and 38.5 %
Oxide and sulphide leach pad capacity limit (Tonnes)	18.2 and 81.8 %
Arsenic grade limit for processing mills (PPM)	1 %

Table 4-6 Operational and economic parameters used at the copper mining complex

Attribute	Value
Number of mining blocks	Mine A: 3600; Mine B: 1200
Production scheduling horizon	13 weeks
Slope angle (Mine A and B)	45, 45
Radius (Mine A and B)	10 mining blocks
Recovery of copper	Oxide leach pad: 65 %; Sulphide leach pad: 27 %; Processing mills: 80.4 %, 80 % and 82.6 %
Recovery of gold, silver, and molybdenum	0.25
Selling cost – processing mills, oxide leach pad, and sulphide leach pad	571, 551, and 551 $/tonne

Selling price – copper, gold, silver, and molybdenum concentrate, copper plate	5511, 35.2×10^6, 4.9×10^5, and 1.3×10^4 \$/tonne
Processing cost – processing mills, oxide leach pad, and, sulphide leach pad	5.79, 5.81, and 1.84 \$/tonne
Crushing cost (Crusher 1, 2, 3, 4, and, 5)	0.58 \$/tonne
Mining cost (Depending on depth)	Mine A: 0.4 - 1.27 ; Mine B: 0.52 - 1.09 (\$/tonne)
Penalty cost: arsenic grade limit, capacity limit at crusher, oxide leach pad, sulphide leach pad, and, processing mills	5 \$/ PPM, 1\$/tonne, 1\$/tonne, 1\$/tonne, and 2\$/tonne

Appendix 4.3

This section presents the continuous updating framework for adapting the short-term production schedule of a mining complex with incoming new information.

Input: Deep neural network (DNN) agent f_θ, Monte Carlo tree search (MCTS), ensemble Kalman filter (ENKF), equipment performance simulation (ESIM) parameters

Output: Updated short-term production plan

Data: Drillhole data $dH_{a,m}^I$, sensor material information l_a^{NI}, mineral value chain $[M, \mathcal{T}_m, \mathcal{S}_m, \mathcal{D}, \mathcal{P}]$, planning horizon N_{week}, historical equipment data eI_e^I, sensor equipment information l_e^{NI}, input state s_1, N_{MCTS}^{Train}, , N_{MCTS}^{Update}

1. Generate $\mathbb{S}_{I,a,m}$ using $dH_{a,m}^I$
2. Generate $\mathbb{S}'_{I,e}$ using eP_e^I
3. $f_\theta \leftarrow$ Self-playRL(NN, MCTS, True, $\mathbb{S}_{I,a,m}$, $\mathbb{S}'_{I,e}$, N_{week}, s_1, N_{MCTS}^{Train})
    ▷ Appendix 4.4
4. if (l_a^{NI}): $\mathbb{S}_{U,a,m} \leftarrow$ ENKF($\mathbb{S}_{I,a,m}$, l_a^{NI})   ▷ Sect. 4.3.3.1
5. if (l_e^{NI}): $\mathbb{S}'_{U,e} \leftarrow$ ESIM(eP_e^I, l_e^{NI})   ▷ Sect. 4.3.3.1
6. if ($l_a^{NI} \parallel l_e^{NI}$) execute in parallel:
    a. Self-playRL(DNN,   MCTS,   False, $\mathbb{S}_{U,a,m}$, $\mathbb{S}'_{U,e}$, N_{week}, s_1, N_{MCTS}^{Update})
    b. Self-playRL(DNN, MCTS, True, $\mathbb{S}_{U,a,m}$, $\mathbb{S}'_{U,e}$, N_{week}, s_1, N_{MCTS}^{Train})

Appendix 4.4

This section presents the self-play reinforcement learning algorithm for training a DNN agent to learn short-term production scheduling in a mining complex and then adapt it with incoming new information.

Input: DNN agent f_θ, MCTS, train, planning horizon N_{week}, input state s_1, N_{MCTS}
Output: Trained DNN agent f_θ
Data: $\mathbb{S}_{I,a,m}$, $\mathbb{S}'_{I,e}$, mineral value chain $[M, \mathcal{T}_m, \mathcal{S}_m, \mathcal{D}, \mathcal{P}]$
1. Initialized f_θ randomly
2. for $w < N_{week}$:
    a. Initialize MCTS with root node $s_t = s_1$
    b. Compute:
       $U_{\mathbb{P}_\mathbb{M},j,w} \leftarrow \max\left(\mathbb{S}'_{I,j}(w)\right),\; L_{\mathbb{P}_\mathbb{M},j,w} \leftarrow \min\left(\mathbb{S}'_{I,j}(w)\right), \forall\, j \in \mathcal{D} \cup \mathcal{P}$
    c. Compute $Ave_{\mathbb{P}_\mathbb{M},j,w} \leftarrow \left(\sum_j E[\mathbb{S}'_{I,j}(w)]\right), \forall\, j \in \mathcal{T}_m, \mathcal{S}_m$
    d. Compute $Ave_{\mathbb{P}_\mathbb{M},M,w} \leftarrow \min\left(Ave_{\mathbb{P}_\mathbb{M},\mathcal{T},w}, Ave_{\mathbb{P}_\mathbb{M},\mathcal{S},w}\right)$
    e. while $Ave_{\mathbb{P}_\mathbb{M},M,w} \geq 0$
       i. for $n \leq N_{MCTS}$ $\triangleright$ Appendix 4.5
          1. $s_L \leftarrow$ MCTS.Selection(s_t, f_θ)
          2. Permissible   actions
             $(x_{b,d,L}) \leftarrow$ MCTS.Expansion(s_L)
          3. $p_n, v_n \leftarrow$ MCTS.Evaluation$(f_\theta(s_L, x_{b,d,L}))$
          4. $z_n \leftarrow$ MCTS.Simulation$(s_L, Ave_{\mathbb{P}_\mathbb{M},M,w})$
          5. MCTS.BackUp(s_L, z_n, v_n)
       ii. $\pi_t, z_t \leftarrow$ MCTS.GetStatistics(s_t)
       iii. $x_{b,d,t} \leftarrow$ MCTS.Play(π_t)
       iv. Compute s_{t+1} with $x_{b,d,t}$
       v. Compute $v_{\mathbb{P}_\mathbb{M},M,t,\mathbb{s}_I} \leftarrow \sum_{d \in \mathcal{D}} v_{\mathbb{P}_\mathbb{M},d,t,\mathbb{s}_I}$ with $x_{b,d,L+1}$, $\forall\, \mathbb{s}_I \in \mathbb{S}'_{I,e}$ $\triangleright$ Eq. 4.6
       vi. $t \leftarrow t + 1$
       vii. $Ave_{\mathbb{P}_\mathbb{M},M,w} \leftarrow Ave_{\mathbb{P}_\mathbb{M},M,w} - v_{\mathbb{P}_\mathbb{M},M,t,\mathbb{s}_I}$
       viii. if train: Train f_θ with (s_t, π_t, z_t) for N_L iterations
    f. if train: Train f_θ with all (s_t, π_t, z_t) for N_G iterations
3. Save production plan and forecasts

Appendix 4.5

This section outlines the Monte Carlo tree search algorithm.

Input: DNN agent f_θ, MCTS parameters

Output: $\pi_t, z_t, x_{b,d,t}$

Data: $\mathbb{S}_{I,a,m}$, $\mathbb{S}'_{I,e}$, mineral value chain $[M, \mathcal{T}_m, \mathcal{S}_m, \mathcal{D}, \mathcal{P}]$, planning horizon $Ave_{\mathbb{P}_\mathbb{M},M,w}$, input state s_t

1. for $n \leq N_{MCTS}$
   a. $s_L \leftarrow$ MCTS.GetLeaf(s_t, f_θ) $\triangleright$ Eq. 4.3 and 4.4
   b. Update state $(v_{a,d,L,\mathbb{s}_I}, v_{a,\mathcal{p},L,\mathbb{s}^{\mathbb{J}}_I}, \forall\, a \in \mathbb{P}_\mathbb{R} \cup \mathbb{P}_\mathbb{D} \cup \mathbb{P}_\mathbb{M}, \mathcal{p} \in \mathcal{P}, d \in \mathcal{D})$ with $x_{b,d,t}, \forall t < L, \mathbb{s}_I \in \mathbb{S}'_{I,e}, \mathbb{s}^{\mathbb{J}}_I \in \mathbb{S}_{I,a,m}$ $\triangleright$ Eqs. 4.5-4.11
   c. Legal actions $(x_{b,d,L}) \leftarrow$ MCTS.LegalAction(s_L) $\triangleright$ Appendix 4.6
   d. MCTS.addnode($x_{b,d,L}$) and initialize $N(s_L, x_{b,d,L}), Q(s_L, x_{b,d,L}),$ $P(s_L, x_{b,d,L}), P(s_L, x_{b,d,L}) = 0$
   e. $p_n, v_n \leftarrow f_\theta\left(v_{a,d,L,\mathbb{s}_I}, v_{a,\mathcal{p},L,\mathbb{s}^{\mathbb{J}}_I}, \mathbb{Z}^{I,\mathbb{s}}_{a'm}\right), \forall\, a \in \mathbb{P}_\mathbb{R} \cup \mathbb{P}_\mathbb{D}, \mathcal{p} \in \mathcal{P}, d \in \mathcal{D}, m \in M$
   f. $r_n \leftarrow$ MCTS.Simulation($s_L, Ave_{\mathbb{P}_\mathbb{M},M,w}$):
   g. $t \leftarrow L$
       i. while $Ave_{\mathbb{P}_\mathbb{M},M,w} > 0$:
           1. $t \leftarrow t + 1$
           2. Legal action $(x_{b,d,t}) \leftarrow$ MCTS.LegalAction$P_\pi(s_{t-1})$
           3. $x_{b,d,t} \leftarrow$ MCTS.RandomSelect$\left((x_{b,d,t})\right)$
           4. $v_{\mathbb{P}_\mathbb{M},M,t,\mathbb{s}_I} \leftarrow \sum_{d \in \mathcal{D}} v_{\mathbb{P}_\mathbb{M},d,t,\mathbb{s}_I}$ with $x_{b,d,t}\ \forall\, \mathbb{s}_I \in \mathbb{S}'_{I,e}$
           5. $Ave_{\mathbb{P}_\mathbb{M},M,w} \leftarrow Ave_{\mathbb{P}_\mathbb{M},M,w} - v_{\mathbb{P}_\mathbb{M},M,t,\mathbb{s}_I}$
       ii. Compute $v_{a,j,t,\mathbb{s}^{\mathbb{J}}_I}, d^+_{a,j,t,\mathbb{s}^{\mathbb{J}}_I}, d^-_{a,i,t,\mathbb{s}^{\mathbb{J}}_I} \forall a \in \mathbb{P}_\mathbb{D} \cup \mathbb{P}_\mathbb{M} \cup \mathbb{P}_\mathbb{R}, j \in \mathcal{P} \cup \mathcal{D} \cup M$ and then r_n $\triangleright$ Eqs. 4.5-4.14
   h. MCTS.BackUp(s_L, r_n, v_n):
       i. Update visit count, reward, and mean reward for all $x_{b,d,t}$ until $\forall t \leq L$ $\triangleright$ Eqs. 4.17-4.19
2. $\pi_t, z_t \leftarrow$ MCTS.GetStatistics(s_t):
   a. Get the MCTS search probability and scalar evaluation for state s_t
3. $x_{b,d,t} \leftarrow$ MCTS.Play(π_t):
   a. $x_{b,d,t} = Multinomial(\pi_t)$

Appendix 4.6

This section outlines the algorithm to find the permissible block extraction and block destination decisions.

Input: S_m, input state s_t
Output: All legal actions $(x_{b,d,t})$ for state s_t
Data: Slope angle (A) and radius (r)
1. for m in M:
    a. for s_i in S_m:
      i. $R \leftarrow 1$
      ii. Find vertical successors blocks $V(s_i, R)$ for s_i
      iii. Find horizontal successors blocks $H(s_i, R)$ for s_i
      iv. for block j in $V(s_i, R)$ and $H(s_i, R)$:
        1. Find vertical predecessor $V(j, A)$ of block j
          a. if all the blocks in $V(j, A)$ is already mined then j is a valid block $\triangleright$ Eq. 4.1
          b. $AcessibleBlock(s_i) \leftarrow j$
        2. $r \leftarrow r - 1$
      v. if $AcessibleBlock(s_i)$ null and $r \neq 0$: $R \leftarrow R + 1$ and go to (ii)
      vi. else:
        1. for each block k in $AcessibleBlock(s_i)$:
          a. Find legal destinations $d(k)$ $\triangleright$ see Sect. 4.2.1
          b. for d in $d(k)$:
            i. $AcessibleBlockDest(s_i) \leftarrow (j, d)$
      vii. Check and remove an elements of $AcessibleBlockDest(s_i)$ if it exist for any other $AcessibleBlockDest(s_j), \forall s_j \neq s_i$
2. Find all the legal actions $(x_{b,d,t})$ such that each action in $(x_{b,d,t})$ is a set consisting of exactly one element from each $AcessibleBlockDest(s_j), \forall s_i \in S_m$

Appendix 4.7

The DNN agent consists of fully connected:

1. Input layer with 3700 neurons that describes the size of the input state

2. Hidden layer 1: 1000 neurons with Xavier initialization (Glorot and Bengio, 2010) and ReLU activation (Nair and Hinton, 2010)

3. Hidden layer 2: 800 neurons with Xavier initialization and ReLU activation

4. Hidden layer 3: 600 neurons with Xavier initialization and ReLU activation

5. Hidden layer 4: 400 neurons with Xavier initialization and ReLU activation

The output from hidden layer 4 is fed to two separate heads for computing policy and value, respectively. The policy head p consist of fully connected:

1. Policy layer 1: 400 neurons with Xavier initialization and ReLU activation

2. Policy layer 2: 300 neurons with Xavier initialization and ReLU activation

3. A fully connected linear layer that outputs a vector of size 256 corresponding to logit probabilities for all actions

The value head v consist of fully connected:

1. Value layer 1: 400 neurons with Xavier initialization and ReLU activation

2. Value layer 2: 300 neurons with Xavier initialization and ReLU activation

3. A fully connected linear layer of dimension 256 with $\tan h$ non-linearity to output a scaler value in the range $[-1,1]$

Appendix 4.8

This section shows the additional results of the case study presented in the manuscript. Figure 4-14 and Figure 4-15 show, respectively, the (a) forecasts of the initial schedule compared to (b) its risk over the updated uncertainties, and (c) the adapted schedule, for the production limit of mills 2 and 3. The initial and adapted schedules respect the capacity limit of mill 2. The initial schedule has a lower utilization in weeks 7, 11, and 13, compared to a higher utilization of the capacity of mill 2 in the adapted schedule. The initial schedule has a lower utilization in periods 5, 9, and 12 compared to a higher utilization in the adapted schedule of the capacity of mill 3. Figure 4-16 and Figure 4-17 show forecasts for the arsenic content limit for

mills 2 and 3, respectively. The arsenic limit is respected in the initial schedule for mills 2 and 3. The updated supply uncertainty of mine A shows a higher concentration of arsenic; therefore, the risk profile of the initial schedule shows violations in the arsenic content limits for mills 2 and 3 in the initial schedule. The DNN agent adapts the schedule to minimize such violations.

Figure 4-14 Mill 2 production limit forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-15 Mill 3 production limit forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-18 and Figure 4-19 show the (a) forecasts of the initial schedule, (b) risk of the initial schedule over the updated uncertainties, and (c) the adapted schedule for the sulphide leach pad and oxide leach pad production limit. The initial and adapted

197

schedules respect the production limit for both leach pads. However, the adapted schedule processes less material at the leach pads compared to the initial schedule, to achieve a blending of materials at the processing mills that complies with the arsenic content limit.

(a)   (b)   (c)

Figure 4-16 Mill 2 arsenic limit forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-20 and Figure 4-21 show the forecasts for the quantity of recovered silver and molybdenum products, respectively. The forecasts of the initial schedule over the updated uncertainties (Figure 4-20(b)) show an increase of 23% in silver production. The adapted schedule shows an increase of 43% in silver production (Figure 4-20(c)). The forecasts of initial schedule over the updated uncertainties show an increase of 32% (Figure 4-21(b)) in Molybdenum production. Molybdenum production is increased by 61% in the adapted schedule (Figure 4-21(c)). Figure 4-22 shows the forecasts for copper cathode plate production with leach pads for the initial and adapted schedule. The forecasts of the initial schedule over the updated uncertainties show an increase of 5% in copper cathode production; however, the DNN agent adapts the plan to reduce copper cathode production by 2%.

198

Figure 4-17 Mill 3 arsenic limit forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-18 Sulphide leach pad production limit forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-19 Oxide leach pad production limit forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-20 Silver production forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-21 Molybdenum production forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-22 Copper cathode production forecasts for (a) initial short-term production schedule, (b) risk profile of initial schedule over the updated uncertainties, and (c) adapted schedule

Figure 4-23 Block extraction sequence in (a) the initial short-term production schedule compared to (b) the adapted short-term production schedule for bench 1

Figure 4-24 Block destination decisions in (a) the initial short-term production schedule compared to (b) the adapted short-term production schedule for bench 1

Figure 4-23 and Figure 4-24 show the block extraction and destination decisions, respectively, for bench 2 of mines A and B for the (a) initial and (b) adapted short-term production schedules. Although the distribution of metals has not changed significantly in this bench, the DNN agent still adapts the schedule to better blend materials to respect the production limits.

201

CHAPTER 5

Updating Geostatistically Simulated Models of Mineral Deposits in Real-time with Incoming New Information using Actor-Critic Reinforcement Learning

The extended ensemble Kalman filter developed in chapter 3 to update the geostatistically simulated models of pertinent properties of mineral deposits does not learn from the incoming new information and does not account for high-order spatial statistics. This chapter develops a new actor-critic reinforcement learning algorithm that learns from the incoming new information to update geostatistically simulated models of pertinent properties of mineral deposits while accounting for high-order spatial statistics.

5.1   Introduction

New information is readily available with conventional and modern smart digital technologies used during production activities in industrial environments. These technologies include advanced sensors and monitoring devices that are used during production activities in mines and oilfields, and in monitoring activities in the fields of hydrogeology, hydrology, meteorology, atmospheric sciences, geomorphology and oceanography. For example, in an industrial mining environment, global positioning systems can locate and monitor the status of the mining fleet in real-time (Chaowasakoo et al., 2014). Built-in control units can monitor the health and utilization of the mining fleet (Koellner et al., 2004). Radiofrequency identification

tags can locate and track (Rosa et al., 2007), and infrared and X-ray sensors can measure the geological properties of the materials mined, hauled, conveyed, processed, and sold (De Jong, 2004; Goetz et al., 2009; Iyakwari et al., 2016; Dalm et al., 2018). The incoming new information is typically used to update the relevant properties of geostatistically simulated models. However, this new information, referred to as "soft data", is partial and noisy, and is therefore uncertain. The soft nature of the new information is attributed to the characteristics of the related sensors that generate indirect measurements compared to, for example, those derived from the analysis of drillhole samples in geochemical laboratories. Assimilating incoming new information in geostatistically simulated models is similar to history matching in petroleum reservoirs (Oliver et al., 2008; Gilman and Ozgen, 2013). History matching entails using production data such as oil production, flow rates, and well pressure, to update geostatistically simulated models of static reservoir properties, such as porosity and permeability, along with dynamic reservoir properties, such as pressure and fluid saturation, to better match the observed production data.

Ensemble Kalman filter (EnKF) is a thoroughly studied and applied method for history matching in petroleum and groundwater reservoirs (Aanonsen et al., 2009; Evensen, 2009; Oliver and Chen, 2011; Xu et al., 2013). Benndorf (2015, 2020) introduced the use of the EnKF for updating geostatistically estimated models of mineral deposits. Yüksel et al. (2016, 2017) used the EnKF method to update geostatistically simulated models of ash content with incoming new information at a coal deposit. Wambeke and Benndorf (2017) proposed a combination of the EnKF method with a forward simulator, while incorporating a connected updating cycle and a local neighborhood technique, to update the geostatistically simulated models with

the sensor data collected from the conveyor belt at a synthetic mining operation. Wambeke and Benndorf (2018) studied the effect of measurement volumes, blending ratios, and sensor precision within the EnKF method. Wambeke et al. (2018) used a forward simulator and Bond's work theorem with the EnKF method to update the geostatistically simulated models of the Bond work index – a geometallurgical property – at the Tropicana gold mine with the processing mill sensor data about throughput, power draw, feed and product size. Other methods for updating geostatistically simulated models of pertinent properties in a mineral deposit include co-simulation with soft data (Journel and Alabert, 1990; Neves et al., 2018) and conditional simulation by successive residuals (Vargas-Guzmán and Dimitrakopoulos, 2002; Jewbali and Dimitrakopoulos, 2011). Methods such as gradual deformation (Hu, 2000), neighborhood algorithm (Sambridge, 1999), evolutionary algorithm (Schulze-Riegert and Ghedan, 2007), maximum a posteriori (Oliver, 1996), Markov chain Monte Carlo (Oliver et al., 1997; Fu et al., 2017), inverse sequential simulation (Xu and Gómez-Hernández, 2015), classification and regression tree algorithm (Gutiérrez-Esparza and Gómez-Hernández, 2017), randomized maximum likelihood (Sarma et al., 2006; Chen and Oliver, 2012; Vo and Durlofsky, 2014), Tau-model (Naraghi and Srinivasan, 2015), Markov mesh model (Panzeri et al., 2016), variants of the ensemble Kalman filter, and co-simulation with soft data (Journel and Alabert, 1990; Mao and Journel, 1999b; Soares et al., 2017) have also been used for updating geostatistically simulated models of pertinent properties of petroleum and groundwater reservoirs. The abovementioned methods update the relevant properties of geostatistically simulated models but do not learn

from the incoming new information. Additionally, they do not account for nor respect high-order spatial statistics while updating the geostatistically simulated models.

Recent developments in history matching include artificial intelligence (AI) algorithms based on supervised machine learning, such as a convolutional neural network (CNN) with principal component analysis (PCA) (Liu et al., 2019) and stepwise CNN-PCA with recurrent neural network (RNN) (Tang et al., 2019). The CNN-PCA method trains a CNN to post-process a given PCA geostatistical model, which involves using a training dataset to learn to minimize the difference between the style and content of the generated post-processed PCA model, and the target style and content calculated from either a training image or an initial geostatistical model. The CNN-PCA-RNN trains an RNN to generate flow simulation results for given geostatistical models of porosity and permeability, and involves using a training dataset to learn to minimize the difference between the predictions of the RNN and the targets generated by a high-fidelity flow simulator. These methods aim to minimize the mismatch between the targets and generated outputs for a given training dataset and, therefore, cannot perform well if the inputs differ greatly from the training dataset.

The work presented herein, inspired by the continuous control algorithm (Lillicrap et al., 2015), proposes a novel self-learning artificial intelligence algorithm that trains agents (typically function approximators) to update geostatistically simulated models of pertinent properties of mineral deposits in real-time with new incoming spatial and temporal information. New incoming temporal information is sensor data collected over time about the properties of materials at different parts of a mining operation. The proposed algorithm uses deep policy gradient reinforcement

205

learning with an actor and a critic agent (in this work, both are convolutional neural networks) to learn about the relationships between incoming new information and geostatistically simulated models. These relationships are defined by high-order spatial statistics. High-order spatial statistics (Dimitrakopoulos et al., 2010; Mustapha and Dimitrakopoulos, 2011; Minniakhmetov et al., 2018; Yao et al., 2018) can capture complex spatial geological characteristics, curvilinear features, geometric relations, and the connectivity of extreme values needed for updating spatially dependent geological phenomena. In the following sections, first the proposed self-learning AI algorithm is detailed. Next, an application at a copper mining operation is explored to illustrate the efficiency and applied aspects of the proposed algorithm in updating geostatistically simulated models of relevant properties of mineral deposits with incoming new information while reproducing spatial patterns and high-order spatial statistics. Conclusions and directions for future research follow.

5.2 Method

Section 5.2.1 details how new information is collected from different day-to-day production activities in a mining operation. Section 5.2.2 provides the details of the proposed self-learning AI algorithm. Section 5.2.3 details the process of using the proposed algorithm in an operating mining environment. A list of notations used in this section is available in Appendix 5.1. The pseudo-code related to the different parts of the proposed algorithm is outlined in Appendix 5.2-Appendix 5.8.

5.2.1 Incoming new information in a mining operation and related notations

In a mining operation, materials extracted from the mine are sent to destinations and flow through processing streams to generate products that are sold to customers.

Let $Z(x)$ be a spatial random field with random variables $Z(x_i)$, representing a property of a mining block at location x_i, with $i = 1, \ldots N$ being the index of the blocks. Initial direct measurements, I, derived from the analysis of exploration drillhole samples in geochemical laboratories is denoted by D^I. A finite set of initial geostatistically simulated models, $\mathbb{S}^I$, is generated using exploration drillhole samples, D^I, that represent realizations, $z^{\$}(x)$, of the spatial random field, and quantify the uncertainty about the spatial property of materials in the mine. The drilling machines, B, located spatially within the mine, drill holes for the blasting of materials. The sensors installed on the drilling machines measure the quality of materials drilled. The new information, NI, generated spatially by the sensors on the drilling machine about the property of blocks, $Z(x)$, within the mine is denoted by $\mathrm{NI}_B(x)$. The blasted materials are then loaded with shovels, S, into trucks, T. The sensors on the shovels measure the quantity, $q^i(x)$, and quality, $\mathrm{NI}_i(x), \forall i \in S$, of a block, $Z(x)$, loaded. The sensors on the truck measure the quantity, $q^i(x)$, and quality, $\mathrm{NI}_i(x), \forall i \in T$, of $Z(x)$ hauled. The incoming new information about the quality of blocks $Z(x)$ located spatially within the mine is herein referred to as "spatial sensor data" and denoted by $\mathcal{S} \in B \cup S \cup T$. The trucks haul the materials to different destinations, D. Let $f^d, \forall d \in D$ represent a function that mimics the transformation of materials at destination d. The sensors at the destinations measure the quality, $\mathrm{NI}_d(q^d), \forall d \in D$, and quantity, q^d, of materials at destination d. The materials from the destinations are transported via conveyor belts, C, to processing streams, P. The conveyor belt analyzer monitors the rate, $q^{d,p,c}$, and quality, $\mathrm{NI}_c(q^{d,p,c}), \forall d \in D, p \in P, c \in C$, of material transported via conveyor c from destination d to processing stream p. The

processing streams generate the products which are sold to customers. Let $f^p, \forall p \in P$ denote the function that mimics the transformation of materials at processing stream p. The sensors at the processing stream measure the quality, $\text{NI}_p(q^p)$, and quantity, $q^p, \forall p \in P$, of products generated. Let $\mathcal{T}$ represent a set that consists of all the components in a mining operation that handle and process the materials and collect sensor data, i.e. $\mathcal{T} = \{D, C, P\}$. Let $\text{NI}_i(q^i)$ and q^i represent the incoming new information collected over time with sensors at component $i \in \mathcal{T}$ about the quality and quantity of related materials, referred to herein as "temporal sensor data". The tracking devices installed on component $i \in \mathcal{T} \cup \mathcal{S}$ of the mining operation, help to locate and track the flow of materials. Let $Track$ represent an operator that can locate and track the flow of materials from the mine to the destinations and subsequently to the processing streams.

5.2.2   A self-learning AI algorithm

The proposed self-learning AI algorithm for updating the geostatistically simulated models of pertinent properties of mineral deposits with incoming new information uses deep deterministic policy gradient (DDPG) reinforcement learning (Lillicrap et al., 2015) with an actor and a critic agent (in this work, both are convolutional neural networks (CNN)). This class of algorithm is called a model-free deep reinforcement learning (DRL) algorithm. A model-free DRL algorithm is used to update simulated models because the blocks that need to be updated (actions taken by actor agent) are visited in a random path, therefore the actions are not strongly correlated over time. Additionally, no planning is required to update geostatistically simulated models. The agents are only learning the dependencies between an action

and its associated reward in terms of high-order spatial statistics and new incoming information. CNNs are used as agents because they can capture spatial patterns in the inputs, which are present in mineral deposits. The actor, μ, and critic, Q, agents (f_{θ^μ}, and f_{θ^Q} parametrized by θ^μ, and θ^Q, respectively) interact with an environment (see Sect. 5.2.2.1 for details of the environment) in discrete timesteps. Herein, a time step t denotes the point at which a block is visited along a random path visiting all blocks, similar to the sequential simulation approach (Deutsch and Journel, 1992; Journel, 1994). At each time step t the actor takes an action $a_t \in \mathbb{R}$ which is to predict the updated property $Z^{\mathbb{s}'}(x)$ of a block located at x based on a state, s_t – which is fully observable – for a given geostatistically simulated model $\mathbb{s} \in \mathbb{S}^I$. The action is executed in an environment, meaning that the block property, $Z^{\mathbb{s}}(x)$, is updated with the taken action a_t to generate an updated geostatistically simulated model $\mathbb{s}' \in \mathbb{S}^U$ and the agent receives a scalar reward r_t and a next state s_{t+1} (see Sect. 5.2.2.1 for calculation of r_t and s_{t+1}). The process continues until all the blocks are visited and updated. The state, action, reward, next state tuple (s_t, a_t, r_t, and s_{t+1} respectively) from all the time steps is stored in a replay memory buffer, R, of finite-sized cache. The replay memory is sampled uniformly at regular intervals to train both the actor and critic agents (see Sect. 5.2.2.3 for the training of actor and critic agents). The algorithm terminates when N_T iterations are reached, and the trained actor agent f_{θ^μ} can be used to update the geostatistically simulated models, $\mathbb{S}^U$, of pertinent properties of mineral deposits with the spatial and temporal new information collected during day-to-day production activities in a mining operation (see Sect. 5.2.3 for using the agent for an operating mining environment).

5.2.2.1 *Environment*

In the context of the present study an environment is a model of the mining operation that encapsulates how materials extracted from mineral deposits are transformed from raw materials to products with day-to-day production activities. The agents interact with the environment by visiting mining blocks in a random path. The environment provides a representation of the mining operation during its interaction with the agents. The representation is called a state, and includes information such as the property of blocks in initial geostatistically simulated models, conditioning data events and geometry, incoming spatial and temporal new information from sensors, conditional variance of the blocks in the initial geostatistically simulated model and error in the incoming spatial and temporal sensor new information, and the model-based predictions. The environment is also responsible for evaluating and using the action a_t taken by the agent to generate a scalar value r_t, called a reward, and a new representation, called the next state s_{t+1}. The calculation of the state, reward, and next state in the environment are detailed in the following sections.

State

The state s_t is comprised of 10 components. The first component is a property of the blocks in the initial geostatistical simulation $s \in \mathbb{S}^{\mathrm{I}}$ including and surrounding the block located at x in consideration at time step t denoted by $Z^s(Neigh_x)$. Here, $Neigh_x$ defines the radius/neighbourhood to consider around a block located at x, and is an input for the algorithm. If there are missing blocks within the neighbourhood $Neigh_x$ of a block located at x they are initialized with a default value of -1. The second is the conditioning data event $D_x^{s'}$ which includes a property of N^s closest

blocks for a block x in the geostatistically simulated model $s' \in \mathbb{S}^U$. s' is formed by first creating a copy of $s \in \mathbb{S}^I$ and then updating the simulated property of blocks that have been visited until time $t - 1$ with the actions taken by the actor agent. This captures the history of actions taken until $t - 1$. The quantity of conditioning data N^s is an input for the algorithm. For example, suppose $Z^s(Neigh_x)$ and x are respectively represented by the grid and black circle in Figure 5-1(a) then a property of the blocks in the updated geostatistical simulation s' which are closest to a block located at x – represented by the ends of black arrows in Figure 5-1(b) – constitutes $D_x^{s'}$.

Figure 5-1 State representation of (a) a property of the blocks in the neighborhood $Neigh_x$ of a block located at x in consideration at time step t in the initial geostatistically simulated model $s \in \mathbb{S}^I$; (b) conditioning data event found in the geostatistically simulated model $s' \in \mathbb{S}^U$ until $t - 1$; and (c) spatial sensor data with its conditioning data event

The third is the incoming spatial new information $\mathcal{S}$ including and surrounding a block located at x, denoted by $NI_{\mathcal{S}}(Neigh_x)$. The absent values in the spatial new data are initialized to -1. The fourth is the conditioning data event $D_x^{\mathcal{S}}$ found in the spatial sensor data which includes the $N^{\mathcal{S}}$ closest spatial sensors' data. For example, suppose $NI_{\mathcal{S}}(Neigh_x)$ and x are respectively represented by the grid and the black

circle in Figure 5-1(c), then the data collected by the spatial sensors which are closest to the block located at x – represented by the ends of black arrows in Figure 5-1(c) – constitutes D_x^S. The quantity of conditioning data N^S is computed based on the density of the spatial sensor data. The white colour in Figure 5-1(c) denotes that no spatial sensor data was available. The fifth component is the conditional variance $E_x^{\mathbb{S}}$ associated with a property of a block located at x in the initial geostatistically simulated model $\mathbb{s} \in \mathbb{S}^{\mathrm{I}}$ in the neighbourhood $Neigh_x$. $E_x^{\mathbb{S}}$ is calculated as:

$$E_x^{\mathbb{S}} = \frac{\gamma_{\mathbb{s}}}{|Neigh_x|} \sum_{i \in Neigh_x} V\big(Z^{\mathbb{s}}(x_i)\big), \forall \mathbb{s} \in \mathbb{S}^{\mathrm{I}} \tag{5.1}$$

where, $V\big(Z^{\mathbb{s}}(x_i)\big)$ is the conditional variance of the initial simulated property of a block x_i and γ_s is an adjustment factor to adjust the magnitude of the conditional variance of the geostatistically simulated models. The sixth is the error in the spatial sensor data $E_x^{\mathrm{NI}_S}$ collected in the neighbourhood $Neigh_x$ of a block located at x. The error with each of the measured spatial sensor data is an input to the algorithm. $E_x^{\mathrm{NI}_S}$ is computed by averaging the error in the available spatial sensor data in the neighbourhood $Neigh_x$ as:

$$E_x^{\mathrm{NI}_S} = \frac{1}{|Neigh_x|} \sum_{i \in Neigh_x} E\big(\mathrm{NI}_S(x_i)\big) \tag{5.2}$$

The seventh is the incoming temporal sensor data, $q^i, \mathrm{NI}_i(q^i), \forall i \in \mathcal{T}$. The eighth component is the errors in the temporal sensor data, $E^{\mathrm{NI}_i}, \forall i \in \mathcal{T}$, and is an input to the algorithm. The ninth is the model-based prediction for the updated geostatistical

simulation $s' \in \mathbb{S}^U$ at the location of the different temporal sensor data (see sections below for details on how the model-based prediction is generated). Calculating the model-based prediction for the updated geostatistical simulation captures the history of the actions taken until time step $t - 1$. If there are no new temporal sensor data, then such data along with their model-based predictions are initialized to -1. The tenth is the geometry of the conditioning data events $D_x^{s'}$ and D_x^{S}, defined by normalized distance vectors $H_x^{s'}$ and H_x^{S}, respectively (the distance of the conditioning data point from a block located at x).

Model-based predictions

Model-based predictions, MP, calculate the values that should have been observed based on geostatistically simulated models at the location where temporal sensor data was collected. The model-based prediction MP_d^s at a destination d for a geostatistical simulation s is calculated with Eq 5.3.

$$MP_d^s = f^d \left(\sum_{i \in T} Track(q_i(x); Z^s(x)) \right) , \forall s \in \mathbb{S}^I, d \in D \qquad (5.3)$$

Equation 5.3 uses the $Track$ operator first to find the quantities of materials, $q_i(x), \forall i \in T$, hauled with trucks $i \in T$ to a destination d then utilizes the function f^d to calculate how the materials are transformed at the destination, and finally uses the geostatistically simulated property $Z^s(x)$ along with the tracking and transformation data to compute MP_d^s. The model-based prediction MP_c^s for a geostatistical simulation s at a conveyor belt c which transports materials from a destination d to a processing stream p is calculated as:

$$\mathrm{MP}_c^\$ = Track(MP_d^\$) \cdot q^{d,p,c} \quad , \forall \mathbb{s} \in \mathbb{S}^\mathrm{I}, d \in D, p \in P, c \in C \qquad (5.4)$$

Equation 5.4 first uses the $Track$ operator to find the quantity of material $q^{d,p,c}$ flowing from a destination d to a processing stream p with conveyor belt c and then utilizes the model-based prediction from Eq. 5.3 along with the tracking information to compute $\mathrm{MP}_c^\$$. The model-based prediction $\mathrm{MP}_p^\$$ at a processing stream p for a geostatistical simulation $\mathbb{s}$ is calculated as follows:

$$\mathrm{MP}_p^\$ = f^p \left(\sum_{c \in C} Track(\mathrm{MP}_c^\$) \right) \quad , \forall \mathbb{s} \in \mathbb{S}^\mathrm{I}, p \in P \qquad (5.5)$$

Equation 5.5 first uses the $Track$ operator to find the quantity of materials fed to a processing stream with different conveyors, then uses the function f^p to find how materials are transformed into products, and finally uses the model-based prediction from Eq. 5.4 along with the tracking and transformation information to compute $\mathrm{MP}_p^\$$.

Action

The state s_t is fed to the actor agent, f_{θ^μ}, to generate an action $a_t \in \mathbb{R}^+$ as shown below:

$$a_t = f_{\theta^\mu}(s_t) + \mathcal{N}_t \qquad (5.6)$$

where, $\mathcal{N}_t$ is the noise added to the action to ensure exploration during the training phase of the proposed AI algorithm. The action a_t is to predict the updated property, $Z^{\mathbb{s}'}(x)$, of a block located at x based on a state s_t in a geostatistical simulation $\mathbb{s} \in \mathbb{S}^\mathrm{I}$.

Next state

The next state s_{t+1} is generated by first replacing $Z^{\mathbb{s}}(x)$ in $Z^{\mathbb{s}}(Neigh_x)$ with the action a_t taken by the actor agent to form $Z^{\mathbb{s}'}(Neigh_x)$, and then generating the model-based prediction with $Z^{\mathbb{s}'}(x)$. The rest of the information in the next state s_{t+1} remains the same as in the state s_t.

Reward

The action a_t taken by the agent in the state s_t is evaluated by the environment to generate a reward r_t. The reward r_t leverages high-order spatial statistics and consists of three parts, as shown below:

$$r_t = r_{\mathbb{s},t} + r_{\mathcal{S},t} + \sum_{i \in \mathcal{T}} r_{i,t} \tag{5.7}$$

The first part, $r_{\mathbb{s},t}$, evaluates the likelihood of an action a_t in the conditional probability distribution function (CPDF), which is generated by searching for replicates of a conditioning data event $D_x^{\mathbb{s}'}$ within a geostatistically simulated model. The second part, $r_{\mathcal{S},t}$, evaluates the likelihood of an action a_t in the CPDF, which is generated by searching for replicates of a conditioning data event $D_x^{\mathcal{S}}$ within the spatial sensor data. The third part, $r_{i,t}, \forall i \in \mathcal{T}$, computes the error between a model-based prediction and temporal sensor data for an action a_t. The CPDF of a given conditioning data event is generated using high-order spatial Legendre moments (Dimitrakopoulos et al., 2010; Mustapha and Dimitrakopoulos, 2011; Yao et al., 2018). High-order spatial Legendre moments capture multi-point spatial statistics and approximates the CPDF of the center node $Z(x)$ for a given data event $D_x^{\mathbb{s}'}$ using Legendre polynomials. The initial geostatistical simulated property $Z^{\mathbb{s}}(Neigh_x)$

215

inside a neighbourhood $Neigh_x$ is searched for all available replicates, $N_\zeta^\$$, of the conditioning data event, $D_x^{\$'}$, defined by a distance vector, $H_x^{\$'}$. Let $\zeta_{\$,i,j}$ denote the values of each node $j \in N^\$$ in the replicate $i \in N_\zeta^\$$. The replicates $\zeta_{\$,i,j}$ of the data event $D_x^{\$'}$ are used to compute the CPDF of the center node $Z(x)$ with Legendre polynomials as:

$$f_\$\big(Z(x)\big|D_x^{\$'}\big) \approx \widetilde{f_\$^W}\big(Z(x)\big|D_x^{\$'}\big) = \frac{\sum_{i \in N_\zeta^\$} X_i\big(Z(x)\big) \cdot \prod_{j \in N^\$} X_i\big(\zeta_{\$,i,j}\big)}{\sum_{i \in N_\zeta^\$} \prod_{j \in N^\$} X_i\big(\zeta_{\$,i,j}\big)} \quad , \forall \$ \in \mathbb{S}^{\mathrm{I}} \quad (5.8)$$

where,

$$X_i(Z(j)) = \sum_{w=0}^{W} \left(w + \frac{1}{2}\right) P_w(\zeta_{\$,i,j}) P_w(Z(j)) \tag{5.9}$$

W is the degree of the Legendre polynomials. $P_w(Z(x))$ is the wth-degree Legendre polynomial for center node $Z(x)$ calculated as follows:

$$P_w(Z(x)) = \frac{1}{2^w w!} \frac{d^w}{dZ(x)^w} [(Z(x)^2 - 1)^w] \tag{5.10}$$

For more details on the computation of CPDF with high-order spatial Legendre moments, readers are referred to Mustapha and Dimitrakopoulos (2011) and Yao et al. (2018). The first part of the reward calculation, $r_{\$,t}$, which defines the likelihood of a_t compared to the initial geostatistically simulated property $Z^\$(x)$ in the CPDF $f_\$\big(Z(x)\big|D_x^{\$'}\big)$ and is calculated as:

$$r_{\mathbb{s},t} = \lambda_{\mathbb{s}} \cdot \left(\widetilde{f_{\mathbb{s}}^{W}}\big(Z(x) = Z^{\mathbb{s}}(x)\big|D_x^{\mathbb{s}'}\big) - \widetilde{f_{\mathbb{s}}^{W}}\big(Z(x) = Z^{\mathbb{s}'}(x)\big|D_x^{\mathbb{s}'}\big) \right), \forall \mathbb{s} \in \mathbb{S}^{\mathrm{I}} \quad (5.11)$$

where,

$$\lambda_{\mathbb{s}} = \frac{1 - E_x^{\mathbb{s}}}{1 - E_x^{\mathbb{s}} + 1 - E_x^{\mathrm{NI}_{\mathcal{S}}} + \sum_{i \in \mathcal{T}} 1 - E^{\mathrm{NI}_i}} \qquad , \forall \mathbb{s} \in \mathbb{S}^{\mathrm{I}} \quad (5.12)$$

$\lambda_{\mathbb{s}}$ is the weight associated with reproducing the spatial statistics of the initial geostatistically simulated model $\mathbb{s}$. The second part of the reward calculation, $r_{\mathcal{S},t}$, which defines the likelihood of action a_t compared to the initial geostatistically simulated property $Z^{\mathbb{s}}(x)$ in the CPDF $f_{\mathcal{S}}\big(Z(x)\big|D_x^{\mathcal{S}}\big)$ is calculated as:

$$r_{\mathcal{S},t} = \lambda_{\mathcal{S}} \cdot \left(\widetilde{f_{\mathcal{S}}^{W}}\big(Z(x) = Z^{\mathbb{s}}(x)\big|D_x^{\mathcal{S}}\big) - \widetilde{f_{\mathcal{S}}^{W}}\big(Z(x) = Z^{\mathbb{s}'}(x)\big|D_x^{\mathcal{S}}\big) \right), \forall \mathbb{s} \in \mathbb{S}^{\mathrm{I}} \quad (5.13)$$

where,

$$f_{\mathcal{S}}\big(Z(x)\big|D_x^{\mathcal{S}}\big) \approx \widetilde{f_{\mathcal{S}}^{W}}\big(Z(x)\big|D_x^{\mathcal{S}}\big) = \frac{\sum_{i \in N_{\zeta}^{\mathcal{S}}} X_i\big(Z(x)\big) \cdot \prod_{j \in N^{\mathcal{S}}} X_i\big(\zeta_{\mathcal{S},i,j}\big)}{\sum_{i \in N_{\zeta}^{\mathcal{S}}} \prod_{j \in N^{\mathcal{S}}} X_i\big(\zeta_{\mathcal{S},i,j}\big)} \qquad (5.14)$$

and

$$\lambda_{\mathcal{S}} = \frac{1 - E_x^{\mathrm{NI}_{\mathcal{S}}}}{1 - E_x^{\mathbb{s}} + 1 - E_x^{\mathcal{S}} + \sum_{i \in \mathcal{T}} 1 - E^{\mathrm{NI}_i}} \qquad , \forall \mathbb{s} \in \mathbb{S}^{\mathrm{I}} \quad (5.15)$$

$\lambda_{\mathcal{S}}$ is the weight associated with reproducing the statistics of the spatial sensor data. $\widetilde{f_{\mathbb{s}}^{W}}\big(Z(x)\big|D_x^{\mathcal{S}}\big)$ is the CPDF and is computed using Eq. 5.14. For this, the spatial sensor data $\mathrm{NI}_{\mathcal{S}}(Neigh_x)$ is searched for all replicates, $N_{\zeta}^{\mathcal{S}}$, of the conditioning data event, $D_x^{\mathcal{S}}$, defined by a distance vector, $H_x^{\mathcal{S}}$. Let $\zeta_{\mathcal{S},i,j}$, denote the values of each node

$j \in N^{\mathcal{S}}$ in the replicate $i \in N_\zeta^{\mathcal{S}}$. The value of the replicates $\zeta_{\mathcal{S},i,j}$ are used in Eqs. 5.9 and 5.10 to compute the CPDF with Eq. 5.14. The third part of the reward calculation, $r_{i,t}, \forall i \in \mathcal{T}$, which defines the difference between the model-based prediction and the temporal sensor data for action a_t and the initial geostatistically simulated property $Z^{\$}(x)$ is calculated as:

$$r_{i,t} = \lambda_i \cdot \left(\left| \mathrm{MP}_i^{\$} - \mathrm{NI}_i \right| - \left| \mathrm{MP}_i^{\$'} - \mathrm{NI}_i \right| \right) \cdot \gamma_i^{\mathrm{MP}} \quad , \forall i \in \mathcal{T}, \$ \in \mathbb{S}^{\mathrm{I}}, \$' \in \mathbb{S}^{\mathrm{U}} \quad (5.16)$$

where,

$$\lambda_i = \frac{1 - E^{\mathrm{NI}_i}}{1 - E_x^{\$} + 1 - E_x^{\mathrm{NI}_{\mathcal{S}}} + \sum_{i \in \mathcal{T}} 1 - E^{\mathrm{NI}_i}} \quad , \forall i \in \mathcal{T}, \$ \in \mathbb{S}^{\mathrm{I}} \quad (5.17)$$

$\mathrm{MP}_i^{\$}$ and $\mathrm{MP}_i^{\$'}$ are the model-based predictions at component $i \in \mathcal{T}$ (calculated using Eqs. 5.3-5.5) for the initial geostatistically simulated property $Z^{\$}(x)$ and the action a_t respectively. $\lambda_i, \forall i \in \mathcal{T}$ is the weight associated with minimizing the difference between the model-based prediction and the temporal sensor data at component i. $r_{i,t}$ is a subtraction of two differences, as seen in Eq. 5.16, as opposed to $r_{\$,t}$ and $r_{\mathcal{S},t}$, which are subtractions of two probabilities. Therefore, an adjustment factor γ_i^{MP} is used in Eq. 5.16 to ensure that the third part of the reward calculation is of the same magnitude as the other parts.

5.2.2.2 *Actor and critic architecture*

The actor agent, f_{θ^μ}, is a CNN which takes as input the state s_t which includes the initial geostatistically simulated model, spatial sensor data, and some additional data. The additional data includes the temporal sensor data, conditional variance of

the initial geostatistically simulated models (Eq. 5.1), error in the spatial (Eq. 5.2) and temporal sensor data, and model-based predictions (Eqs. 5.3-5.5). The additional data are not added as inputs until the fully connected layer in the actor agent, as shown in Figure 5-2(a). The actor agent takes an action a_t based on the state s_t, which is to predict the updated geostatistically simulated property $Z^{\mathbb{s}'}(x)$ of a block located at x. The critic agent, $f_{\theta Q}$, is also a CNN, which takes as input the state s_t that includes all the inputs for the actor agent and as an additional input the action taken by the actor agent. Similar to the actor agent, the additional data and the action are not added as inputs until the fully connected layer in the critic agent as shown in Figure 5-2(b). The critic agent is an action-value function which evaluates the action a_t taken by the actor agent in the state s_t.

Figure 5-2 (a) Actor and (b) critic agent configuration in the proposed AI algorithm

5.2.2.3 *Actor-critic training*

The actor and critic agents are trained using DDPG reinforcement learning, as shown in Figure 5-3. The actor and critic agents are initialized randomly at time, $t = 1$, with weights, θ^{μ} and θ^{Q}, respectively. In addition to the actor and critic agents, two target agents (target actor and target critic), denoted by $f'_{\theta^{\mu}}$, and $f'_{\theta Q}$, and

parametrized by $\theta^{\mu'}$ and $\theta^{Q'}$, are created to avoid divergence issues. The parameters of the target network are initialized to the actor and critic parameters as $\theta^{\mu'} \leftarrow \theta^{\mu}$ and $\theta^{Q'} \leftarrow \theta^{Q}$. The replay memory buffer is initialized at time, $t = 1$.

Figure 5-3 Actor and critic agents training in the proposed AI algorithm

A random path is defined to visit all the blocks in the mineral deposit, and the point at which a block located at x is visited along this path is referred to as a time step t. At every time step t an action is taken by the actor agent based on the state s_t generated from the environment (see Sect. 5.2.2.1) for the mining block in consideration. The action is evaluated in the environment to generate the reward r_t and the next state s_{t+1}. The state, action, reward, and next state tuple (s_t, a_t, r_t, and s_{t+1} respectively) is stored in replay memory R. At every N_I iteration the memory is

sampled to generate mini batches of transitions (s_t, a_t, r_t, and s_{t+1}) of size N_{BS}. The sampled mini batches are used to train the actor and critic agents. More specifically, the parameters, θ^Q, of the critic agent are updated to minimize the temporal difference error loss L given by:

$$L = \frac{1}{N_{BS}} \sum_{i \in N_{BS}} \left(r_i + \gamma . f'_{\theta^Q}\left(s_{i+1}, f'_{\theta^\mu}(s_{i+1})\right) - f_{\theta^Q}(s_i, a_i) + c \cdot \|\theta^Q\|^2 \right) \quad (5.18)$$

$c \cdot \|\theta^Q\|^2$ is an L2 regularization added to the loss function with a penalty cost of c, to avoid overfitting. The actor agent is trained by the sampled policy gradient given as:

$$\nabla_{\theta^\mu} J \approx \frac{1}{N_{BS}} \sum_{i \in N_{BS}} \left(\nabla_{f_{\theta^\mu}(s_i)} f_{\theta^Q}\left(s_i, f_{\theta^\mu}(s_i)\right) \nabla_{f_{\theta^\mu}(s_i)} f_{\theta^\mu}(s_i) \right) \quad (5.19)$$

The sampled policy gradient first takes the gradient of the critic agent parameters, θ^Q, with respect to the action a_t taken by the actor, and then takes the gradient of the actor agent parameters, θ^μ, with respect to the action a_t. The parameters of the trained actor and critic agents are then used to perform soft updates to the target agents as follows:

$$\theta^{\mu'} \leftarrow \tau\theta^\mu + (1 - \tau)\theta^{\mu'} \quad (5.20)$$

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1 - \tau)\theta^{Q'} \quad (5.21)$$

τ defines the strategy to blend the target agent parameters with the trained agent parameters. The new parameters of the actor and critic agents are used for further learning.

5.2.3 Responding to incoming new information

The proposed AI algorithm in Sects. 5.2.2.1-5.2.2.3 trains the CNN actor agent which can update the geostatistically simulated models of pertinent spatial properties of mineral deposits with new spatial and temporal information collected in an operating mining environment.

Figure 5-4 Real-time learning and updating with incoming new spatial and temporal information

The spatial sensor data, $\mathrm{NI}_i(x), i \in \mathcal{S}$, and temporal sensor data, $q^i, \mathrm{NI}_i(q^i)$, $\forall i \in \mathcal{T}$, along with the tracking sensor data, $Track$, and initial geostatistically simulated models of pertinent spatial properties of the mineral deposit, $s \in \mathbb{S}^{\mathrm{I}}$, are fed to the AI

algorithm as shown in Figure 5-4. The AI algorithm then initializes the environment presented in Sect. 5.2.2.1 with all the information. A random path is then defined by the environment to visit all the blocks within the mineral deposit. At each block, a state generated by the environment is fed to the trained actor agent that takes an action, which is to predict the updated geostatistically simulated property of the block. The environment uses the action to generate the next state, and the process continues until all the blocks in the random path are visited. The updated geostatistically simulated property of all the blocks forms the set of updated geostatistically simulated models $\mathbb{S}^U$. The updated models are then used to generate the updated model-based predictions. In parallel, the actor and critic agent parameters are updated by training over the updated models and newly collected information, as presented in Sects. 5.2.2.1-5.2.2.3.

5.3   Application at a synthetic copper mining operation

The proposed self-learning AI algorithm is programmed using Python and Tensorflow. It is applied in this section to a fully known public dataset (Mao and Journel, 1999a), which represents a copper deposit in this case study. 20 initial geostatistically simulated models of copper grades are generated with 416 drillhole data points sampled from a section of the fully known dataset using random stratified sampling. The incoming spatial sensor data of copper grades is generated by randomly sampling 2600 data points from the same sections of the fully known dataset. The error in each spatial sensor data point is sampled from a normal distribution with a mean of 0 and a standard deviation of 0.45. The temporal sensor data of copper grades from the processing mill is generated by randomly sampling the same section of the

fully known dataset in such a way that it imitates the process of collection of the temporal sensor data in a mining operation. The generated temporal sensor data contains 2600 data points. The error in each temporal sensor data point is sampled from a normal distribution with a mean of 0 and a standard deviation of 0.6. The standard deviation of the normal distribution for error in the temporal sensor data is lower than in spatial sensor data to reflect the quality of the respective sensors.

Figure 5-5 The copper mining operation

The copper mining operation considered in this case study is shown in Figure 5-5, and consists of a mine, a waste dump, a processing mill, and a customer. Multiple drilling machines located at the mine perform the drilling operations and capture the spatial new information about the grade of the drilled mining blocks. The materials from the mine are extracted by two shovels and are loaded into trucks which haul the materials to either a processing mill or a waste dump. The processing mill blends and processes the received materials to generate copper products, which are transported to the customers. The sensors at the processing mill capture the temporal new information by measuring the grade of the generated copper products.

Two different datasets – training and testing – are generated using the process outlined above to represent the operations and collection of new information in two areas, Area 1 and Area 2, of the deposit. Each area of the deposit consists of 13,000 mining blocks. The data from Area 1 of the deposit is used to train the proposed AI algorithm. In an operating mine, this dataset would be the historical data. Section 5.3.1 discusses the training results and shows the learning capabilities of the proposed algorithm. The data from Area 2 of the deposit is used for testing the proposed algorithm. This dataset was never used for training and represents how the proposed algorithm would be used in an operating mine to update the geostatistically simulated models of pertinent properties of the deposit with incoming new information. Section 5.3.2 discusses the testing results of the proposed algorithm and shows that the algorithm can generalize and has practical applications in an operating mine. Throughout the presentation and discussion of the results, the fully known data and its histogram, variogram and spatial cumulant maps are also shown for reference purposes. The fully known data is referred to hereafter as the ground truth model. The algorithm takes less than 30 seconds to update the copper grade simulated models of the different areas of the deposit. The architecture of the agents and the value of the different parameters used in this case study are outlined in Appendix 5.9 and Appendix 5.10, respectively.

5.3.1 Training dataset

This section presents the results of training the proposed AI algorithm at Area 1 of the deposit and highlights its learning capabilities. Figure 5-6 shows the drillhole data and two of the initial simulated copper grade models for Area 1 of the deposit.

The initial models in Figure 5-6 show the presence of vertical curvilinear structures in Area 1 of the deposit.

Figure 5-6 Drillhole data and the 2 of the initial copper grade simulated models for Area 1 of the deposit

Figure 5-7 (a) Spatial sensor data; (b) error in the spatial sensor data; and (c) processing mill sensor data of copper grades from Area 1 of the deposit

226

Figure 5-7(a) shows the spatial sensor data collected during the drilling operations in Area 1 of the deposit. The spatial sensor data shows the presence of a similar vertical curvilinear structure at this area of the deposit, however, Figure 5-7(b) shows the error in the spatial sensor data and thereby introduces uncertainty about the visual assessment. Figure 5-7(c) shows the processing mill sensor data collected during operations in Area 1 of the deposit. The black line represents the actual measurement, and the gray lines show the error in the processing mill sensor data.

Figure 5-8 (a) One of the initial simulated models compared to (b) its corresponding updated model and (c) the ground truth model of copper grades for Area 1 of the deposit

Figure 5-8(a) and (b) show one of the initial and its corresponding updated copper grade simulated models for Area 1 of the deposit. The updated model shows more connected high-grade vertical curvilinear structures inferred from the noisy spatial sensor data when compared to the initial model, while still maintaining the geological structures found in the initial model. Figure 5-8(c) shows the ground truth model of

227

copper grades for Area 1 of the deposit for reference only. The algorithm does not have access to the ground truth model. The updated model closely resembles the structures and geological patterns seen in the ground truth model.

Figure 5-9 Histogram and variogram of (a-b) the initial simulated models compared to (c-d) the updated models of copper grades for Area 1 of the deposit

Figure 5-10 Third- and fourth-order spatial cumulant maps for (a-b) the initial simulated model compared to (c-d) its corresponding updated model and (e-f) the ground truth model of copper grades for Area 1 of the deposit

Figure 5-9(a) and (b) show the histogram and variogram, respectively, for the initial copper grade simulated models for Area 1 of the deposit. The initial models reproduce the histogram and variogram of drillhole data. The updated models have a somewhat different distribution of materials (histogram in Figure 5-9(c)) compared to the initial models. The histogram of the updated models reproduces the histogram of the drillhole data and more closely resembles the histogram of the spatial sensor data. The variogram of the updated models, as shown in Figure 5-9(d), reproduces the variogram of the drillhole data and resembles the variogram of the spatial sensor data. Figure 5-10(a) and (b) show the 3rd and 4th order spatial statistics (cumulant), respectively, for the initial copper grade simulated model of Area 1 of the deposit. Figure 5-10(c) and (d) show the 3rd and 4th order spatial cumulant maps, respectively, for the corresponding updated model of Area 1 of the deposit. The 3rd and 4th order

spatial cumulant maps for the updated model show the presence of more connected structures, which are also seen in the spatial cumulant maps of the ground truth model (Figure 5-10(e-f)).

Figure 5-11 Model-based prediction generated with (a) an initial simulated model compared to (b) its corresponding updated model of copper grades for Area 1 of the deposit

Figure 5-11(a) shows the model-based prediction generated with an initial copper grade simulated model of Area 1 of the deposit, for the collected processing mill sensor data. The model-based prediction generated with an initial model shows significant deviations from the processing mill sensor measurements. However, the model-based prediction generated with its corresponding updated model, as shown in Figure 5-11(b), shows minimal deviations from the processing mill sensor measurements. The model-based prediction generated with the updated model still has some deviations because the proposed AI algorithm is not trying and should not

get as close as possible to the sensor measurements, given that the sensor measurements are soft, therefore uncertain.

5.3.2 Testing dataset

The proposed AI algorithm is applied at Area 2 of the deposit, which was not used to train the algorithm.

Figure 5-12 Drillhole data and two of the initial copper grade simulated models for Area 2 of the deposit

Figure 5-12 shows the drillhole data and two of the initial copper grade simulated models for Area 2 of the deposit. The initial models for Area 2 in Figure 5-12 show a presence of very different geological patterns compared to Area 1: the curvilinear structures are horizontal instead of vertical. Figure 5-13(a-c) show the spatial sensor data, the error in the spatial sensor data and the processing mill sensor data, respectively, collected during operations in Area 2 of the deposit.

Figure 5-13 (a) Spatial sensor data; (b) error in the spatial sensor data; and (c) processing mill sensor data of copper grades from Area 2 of the deposit

Figure 5-14 (a) One of the initial simulated models compared to (b) its corresponding updated model and (c) the ground truth model of copper grades for Area 2 of the deposit

Figure 5-14(a-b) show one of the initial and its corresponding updated simulated models of copper grades, respectively, for Area 2 of the deposit. The proposed AI

algorithm can reproduce the horizontal curvilinear structures in the updated model as inferred from the initial model and the spatial sensor data for this area, even though the training data set had vertical curvilinear structures. In addition, the updated model closely resembles the ground truth model shown in Figure 5-14(c).

Figure 5-15 Histogram and variogram of (a-b) the initial simulated models compared to (c-d) the updated models of copper grades for Area 2 of the deposit

Figure 5-16 Third- and fourth-order spatial cumulant maps for (a-b) the initial simulated model compared to (c-d) its corresponding updated model and (e-f) the ground truth model of copper grades for Area 2 of the deposit

Figure 5-17 Model-based prediction generated with (a) an initial simulated model compared to (b) its corresponding updated model of copper grades for Area 2 of the deposit

234

The initial and updated models are validated through the analysis of the histogram and variogram reproductions in Figure 5-15. The initial and updated models both respect the histogram and variogram of the drillhole data. The updated models present a closer resemblance to the histogram and variogram of the spatial sensor data, as shown in Figure 5-15(c) and Figure 5-15(d), respectively. Figure 5-16(a-b) show the 3rd and 4th order spatial cumulant maps, respectively, for the initial copper grade simulated model for Area 2 of the deposit. The spatial cumulant maps of the updated model (Figure 5-16(c-d)) show more connected structures, as seen in the spatial cumulant maps of the ground truth model (Figure 5-16(e-f)). The proposed AI algorithm was not trained with data from this area, yet still reproduces the histogram, variogram, and spatial cumulant maps while updating the initial geostatistically simulated models of copper grades for Area 2 of the deposit. Figure 5-17(a-b) show the model-based prediction for an initial and its corresponding updated simulated model of copper grades, respectively, for Area 2 of the deposit. The model-based prediction for the updated model shows fewer deviations from the processing mill sensor measurements when compared to the initial model.

5.4   Conclusions

This paper proposes a new self-learning artificial intelligence algorithm that uses deep policy gradient reinforcement learning and leverages high-order spatial statistics to train actor and critic agents to update the simulated models of pertinent spatial properties of a mineral deposit with incoming new information. The algorithm is general and can be applied to any mining operation with multiple sources of incoming spatial and temporal new information. The algorithm visits the mining blocks within

235

the mineral deposit following a random path. For each mining block, a state representation is generated and fed to the actor agent. The actor agent, in this case, a convolutional neural network, takes an action, which is to predict the updated spatial properties of blocks based on the state representation. The action is evaluated by the critic agent, in this case, also a convolutional neural network. In parallel, the environment also evaluates the action using high-order spatial statistics and generates a reward and the next state representation. The state, action, reward, next state data is stored in a replay memory, which is sampled at regular intervals to train the agents. The improved agents are then used for further training. An application of the proposed algorithm at a synthetic copper mining operation demonstrates its efficiency and applied aspects. The case study shows that the algorithm can account for softness in the incoming new information (whether spatial or temporal) to update the copper grade simulated models of the deposit while reproducing geological patterns and high-order spatial statistics. The case study also highlights the learning and generalization capabilities of the algorithm by its application in different parts of the deposit, which have different geological patterns and curvilinear structures. Future research will focus on expanding and applying the algorithm for multiple elements.

Appendix 5.1

This section presents the notations used in the proposed self-learning AI algorithm. Table 5-1 shows the sets, indices, and constants used in the proposed algorithm. Table 5-2 presents the variables used in the proposed algorithm.

Table 5-1 Sets, indices, and constants used in the proposed algorithm

Parameters	Definition
$Z(x)$	Spatial random field consisting of random variables $Z(x_i)$, $\forall i \in [1, N]$
$Z(x_i)$	Random variable representing a property of a mining block at location x_i, $\forall i \in [1, N]$ within the mine
D^I	Initial (I) drillhole (D) samples at the mine
$\mathbb{S}^I$	Set of initial stochastic simulations generated with D^I for all blocks $z(x) \in Z(x)$ in the mine
$Z^{\mathbb{s}}(x)$	A property of a block located at x in the initial geostatistical simulation $\mathbb{s} \in \mathbb{S}^I$
$\mathbb{S}^U$	Set of updated geostatistically simulated models for all blocks within the mine; $\mathbb{s}' \in \mathbb{S}^U$
$Z^{\mathbb{s}'}(x)$	A property of a block located at x in the updated geostatistical simulation $\mathbb{s}' \in \mathbb{S}^U$
B, T, S, D, C, P	Set of blasthole drilling machines, trucks, shovels, destinations, conveyor belts, and processing streams, respectively in a mining operation
$\text{NI}_B(x)$	New spatial information (NI) generated by sensors in blastholes (B) about the quality of material drilled
$q^i(x), \text{NI}_i(x)$	New information about the quantity and quality of materials, respectively at component $i \in S \cup T$
$\mathcal{S}$	Set of sensors that generate spatial new information, i.e. $\mathcal{S} \in B \cup S \cup T$
$q^d, \text{NI}_d(q^d)$	Temporal new information generated by the sensor at destination $d \in D$ about the quantity and quality of materials, respectively
f^d	A function that mimics the transformation of materials at destination $d \in D$

$q^{d,p,c}, \mathrm{NI}_c(q^{d,p,c})$	Temporal new information generated by the sensor at conveyor belt $c \in C$ about the quantity and quality of materials, respectively transported from destination $d \in D$ to processing stream $p \in P$
$q^p, \mathrm{NI}_p(q^p)$	Temporal new information generated by the sensor at processing stream $p \in P$ about the quantity and quality of materials, respectively
f^p	A function that mimics the transformation of materials at processing stream $p \in P$
$\mathcal{T}$	Set of sensors that generate temporal new information, i.e. $\mathcal{T} \in D \cup C \cup P$
$Track$	An operator that tracks the flow of materials in a mining operation
$Neigh_x$	The neighbourhood of mining block at location x that also includes block located at x
$Z^{\mathbb{s}}(Neigh_x)$	Simulated value of all the block in simulation $\mathbb{s}$ in the neighbourhood $Neigh_x$
$D_x^{\mathbb{s}'}$	Conditioning data event value for a block located at x in updated simulation $\mathbb{s}'$
$N^{\mathbb{s}}$	Number of conditioning values in the data event in the simulation
$N_\zeta^{\mathbb{s}}$	Number of replicates for data event $D_x^{\mathbb{s}'}$ found in the initial simulation $\mathbb{s}$ within a neighbourhood $Z^{\mathbb{s}}(Neigh_x)$
$\zeta_{\mathbb{s},i,j}$	Value of node $j \in N^{\mathbb{s}'}$ in the replicate $i \in N_\zeta^{\mathbb{s}}$
$\mathrm{NI}_S(Neigh_x)$	Spatial sensor data in the neighbourhood $Neigh_x$
$D_x^{\mathcal{S}}$	Conditioning data event value for block located at x in spatial sensor data
$N^{\mathcal{S}}$	Number of conditioning values in the data event in the spatial sensor data
$N_\zeta^{\mathcal{S}}$	Number of replicates for data event $D_x^{\mathcal{S}}$ found in the spatial sensor data $\mathcal{S}$ within a neighborhood $\mathrm{NI}_S(Neigh_x)$
$\zeta_{\mathcal{S},i,j}$	Value of node $j \in N^{\mathcal{S}}$ in the replicate $i \in N_\zeta^{\mathcal{S}}$
$V(z^{\mathbb{s}}(x_i))$	The variance of a block at a location x_i computed over the set of simulations $\mathbb{S}^I$
$E_x^{\mathbb{s}}$	Average conditional variance of the initial simulations $\mathbb{s}$ in neighbourhood $Neigh_x$ of a block located at x
$\gamma_{\mathbb{s}}$	Adjustment factor to adjust the magnitude of $E_x^{\mathbb{s}}$

$E\left(\mathrm{NI}_{\mathcal{S}}(x_i)\right)$	The error of spatial sensor data at the location of the mining block x_i
$E_x^{\mathrm{NI}_{\mathcal{S}}}$	Average error with the spatial sensor data in the neighborhood $Neigh_x$ of a block located at x
E^{NI_i}	Error with new information collected from the component $i \in D \cup C \cup P$ in a mining operation
$H_x^{\mathbb{s}'}$	The geometry of conditioning data event $D_{x_c}^{\mathbb{s}'}$ found in the initial simulation defined by a normalized three-dimensional distance vector
W	Legendre series polynomial order for the approximation of CPDF
H_x^S	The geometry of conditioning data event $D_{x_c}^B$ found in the spatial sensor data defined by a normalized three-dimensional distance vector
$Z^{\mathbb{s}'}(Neigh_x)$	Simulated value of all the block in the updated simulation $\mathbb{s}'$ in the neighbourhood $Neigh_x$
τ	The soft target update parameter
N_{BS}	Batch size
γ	Discount factor
N_R	Replay memory cache size
N_I	Training interval
γ_i^{MP}	The adjustment factor for adjusting the magnitude of model-based prediction error
c	L2 regularization cost
N_U	Update training iterations
N_{TE}	Number of training episodes
N_T	Number of training iterations
N_{UE}	Number of update training episodes

Table 5-2 Variables used in the proposed algorithm

Variable	Definition
f_{θ^μ}	Actor CNN agent parameterized by θ^μ
f_{θ^Q}	Critic CNN agent parametrized by θ^Q
s_t	State at time t
a_t	Action proposed by actor agent at time t
r_t	Reward computed by the environment at time t
$\mathcal{N}_t$	Random noise process added to the actions at time t, for exploration during training

f'_{θ^Q}	Target critic CNN agent parameterized by $\theta^{Q'}$
f'_{θ^μ}	Target actor CNN agent parameterized by $\theta^{\mu'}$

Appendix 5.2

This section outlines how to use the algorithm in an operating mining environment to update the geostatistically simulated models of pertinent spatial properties of the mineral deposit with incoming spatial and temporal new information.

<u>Input</u>: Actor and critic parameters (f_{θ^μ} and f_{θ^Q}), Environment (Env), training iterations (N_T), update training iterations (N_U), training episodes (N_{TE}), update training episodes (N_{UE})

<u>Data</u>: Drillhole data (D^I), incoming new information ($NI_i(x), i \in S; q^i, NI_i(q^i), \forall i \in \mathcal{T}$), material flow sensor data ($Track$)

<u>Output</u>: Updated stochastic simulations ($\mathbb{S}^U$), trained actor and critic agents ($f_{\theta^\mu}, f_{\theta^Q}$), updated model-based prediction ($MP_i^{\mathbb{s}'}, \forall i \in \mathcal{T}, \mathbb{s}' \in \mathbb{S}^U$)

1. Generate initial stochastic simulations $\mathbb{S}^I$ using D^I
2. Generate historical spatial sensor data $H_i(x), \forall i \in S$, temporal sensor data $q^i, H_i(q^i), \forall i \in \mathcal{T}$, and tracking sensor data ($HTrack$) for training the agents
3. Initialize environment: $Env \leftarrow Env.setup(q^i, H_i(q^i), H_i(x), \mathbb{S}^I, HTrack)$
4. $f_{\theta^\mu}, f_{\theta^Q} \leftarrow \text{ProposedAIAlgorithm}(\theta^\mu, \theta^Q, Train, Env, N_T, N_{TE})$
5. If new information is available, then execute in parallel:
   a. $UpdatedEnv \leftarrow Env.setup(NI_i(x), q^i, NI_i(q^i), \mathbb{S}^I, Track)$
   b. $\mathbb{S}^U \leftarrow \text{ProposedAIAlgorithm}(f_{\theta^\mu}, f_{\theta^Q}, Test, UpdatedEnv)$
   c. $MP_i^{\mathbb{s}'} \leftarrow$
      $UpdatedEnv.ModelPrediction\left(Track, \mathbb{s}', j, \mathbb{S}^U q^i, NI_i(q^i)\right), \forall i \in \mathcal{T}, j \in [1, |Z(x)|], \mathbb{s}' \in \mathbb{S}^U$
   d. $f_{\theta^\mu}, f_{\theta^Q} \leftarrow \text{ProposedAIAlgorithm}(\theta^\mu, \theta^Q, Train, UpdatedEnv, N_T, N_{TU})$

Appendix 5.3

This section presents the deep deterministic policy gradient reinforcement learning algorithm for training the agents.

Input: Actor and critic parameters (f_{θ^μ} and f_{θ^Q}), initial simulation ($\mathbb{S}^I$),Training ($Train$), Environment (Env), training iterations (N_T), episodes (N_E)

Data: Replay memory size (N_R), training interval (N_I), mini-batch size (N_{BS}), soft updated parameter (τ), discount factor (γ), L2 regularization cost (c)

Output: Trained actor and critic agent f_{θ^μ} and f_{θ^Q}

1. Initialized f_{θ^μ} and f_{θ^Q} with θ^μ and θ^Q and replay memory R of size N_R

2. Initialize target actor (f'_{θ^μ}) and critic (f'_{θ^Q}) agent with $\theta^{\mu'} \leftarrow \theta^\mu$ and $\theta^{Q'} \leftarrow \theta^Q$

3. for e in $1..N_E$ do

   a. $nI = 0$

   b. for $\mathbb{s}$ in $\mathbb{S}^I$ do

      i. Initialize $\mathbb{S}^U \leftarrow \mathbb{S}^I$

      ii. Define random path, $Rand([1, |Z(x)|])$ to visit all mining blocks

      iii. if $Train$: Initialize random process $\mathcal{N}$ for action exploration

      iv. for t in $Rand([1, |Z(x)|])$ do $\triangleright$ Here t represent the location of a block

          1. Receive state s_t, $Neigh_t = Env.state(t, \mathbb{s}, \mathbb{S}^U)$

          2. if $Train$:

             Select action $a_t = f_{\theta^\mu}(s_t) + \mathcal{N}_t$

           else:

               Select action $a_t = f_{\theta^\mu}(s_t)$

          3. $Z^{\mathbb{s}'}(t) \leftarrow a_t$; $\forall \mathbb{s}' \in \mathbb{S}^U$

          4. $r_t, MP_{t+1}^{\mathbb{s}'} = Env.reward(\mathbb{s}, s_t, a_t, Neigh_t, Z^{\mathbb{s}}(t), \mathbb{S}^I, \mathbb{S}^U)$

          5. $s_{t+1} = Env.nextstate(s_t, t, a_t, MP_{t+1}^{\mathbb{s}'})$

          6. if $Train$: Store tuple (s_t, a_t, r_t, and s_{t+1}) in R ; and $nI \leftarrow nI + 1$

          7. if $nI \bmod N_I = 0$ and $Train$: for nT in $1..N_T$ do

            a. Sample minibatch of size N_R from R

            b. Update critic by minimizing TD error $L = \frac{1}{N_R}\Sigma_{i\in N_R}\left(\left(r_i + \gamma \cdot f'_{\theta^Q}\left(s_{i+1}, f'_{\theta^\mu}(s_{i+1})\right)\right) - \left(f_{\theta^Q}(s_i, a_i)\right) + c \cdot \|\theta^Q\|^2\right)$

            c. Update actor by sampled policy gradient $\nabla_{\theta^\mu} J \approx \frac{1}{N_R}\Sigma_{i\in N_R}\left(\nabla_{f_{\theta^\mu}(s_i)}f_{\theta^Q}\left(s_i, f_{\theta^\mu}(s_i)\right)\nabla_{f_{\theta^\mu}(s_i)}f_{\theta^\mu}(s_i)\right)$

            d. Update target actor and target critic $\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}$ and $\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$

Appendix 5.4

This section outlines the algorithm for the calculation of model-based predictions.

Input: Material flow operator ($Track$), simulation number ($\mathbb{s}$), block location x,
  simulations ($\mathbb{S}^I$), new information at component i $(q_i, NI_i(q_i))$
Output: $MP_i^{\mathbb{s}}$
1. $MP_i^{\mathbb{s}} \leftarrow []$ ▷ Initialize model-based prediction vector
2. If new information $(q_i, NI_i(q_i))$ is available about mining block x:
   1. $Loc \leftarrow Track(NI_i)$ ▷ Tracking the blocks that generated the sensor data
   2. $MP_i^{\mathbb{s}}.append\left(f^i\left(Track(Z^{\mathbb{s}}(Loc), q_i)\right)\right), \forall \mathbb{s} \in \mathbb{S}^I, i \in \mathcal{T}$ ▷ Eq. 5.3-5.5
3. Else:
   1. $MP_i^{\mathbb{s}}.append(-1)$ ▷ Returning the default value

Appendix 5.5

This section presents the algorithm for the calculations of the next state.

Input: Initial state (s_t) , location (x), action (a_t) , model prediction $MP_{t+1}^{\mathbb{s}'}$
Output: s_{t+1}
1. Initialize $s_{t+1} \leftarrow s_t$
2. $s_{t+1}.SimState[x] \leftarrow a_t$
3. $s_{t+1}.AdditionalDataMP[i] \leftarrow MP_{t+1}^{\mathbb{s}'}[i], \forall i \in \mathcal{T}$

Appendix 5.6

This section presents the algorithm for reward calculations.

Input: Simulation number ($\mathbb{S}$), state (s_t), updated simulation value $\left(Z^{\mathbb{S}'}(x)\right)$, neighborhood ($Neigh_x$), initial simulation value $\left(Z^{\mathbb{S}}(x)\right)$, initial simulations ($\mathbb{S}^{\mathrm{I}}$), updated simulations ($\mathbb{S}^{\mathrm{U}}$)

Data: Degree polynomials (W), temporal new information error adjustment factor (γ_i^{MP})

Output: Reward and new model prediction

1. Reward $\leftarrow 0$, MP $\leftarrow []$

2. $Z^{\mathbb{S}}(Neigh_x) \leftarrow s_t.\text{SimState}, E_x^{\mathbb{S}} \leftarrow s_t.\text{AdditionalDataES}$

3. $D_x^{\mathbb{S}'}, H_x^{\mathbb{S}'} \leftarrow s_t.\text{AdditionalDataSim}, D_x^{\mathcal{S}}, H_x^{\mathcal{S}} \leftarrow s_t.\text{AdditionalDataSpatial}$

4. $\text{NI}_{\mathcal{S}}(Neigh_x) \leftarrow s_t.\text{SpatialState}, E_x^{\text{NI}_{\mathcal{S}}} \leftarrow s_t.\text{AdditionalDataEB}$

5. $E^{\text{NI}_i} \leftarrow s_t.\text{AdditionalDataEI}[i], \text{NI}_i(q_i), q_i \leftarrow s_t.\text{AdditionalDataOS}[i] \; ; \forall\, i \in \mathcal{T}$

6. $N_\zeta^{\mathbb{S}}, \zeta_{\mathbb{S},i,j} \leftarrow Env.\,Replicate(Z^{\mathbb{S}}(Neigh_x), D_x^{\mathbb{S}'}, H_x^{\mathbb{S}'})$

7. $\widetilde{f_{\mathbb{S}}^{W}}\left(Z(x)\big|D_x^{\mathbb{S}'}\right) \leftarrow Env.\,BuildPDF(N_\zeta^{\mathbb{S}}, \zeta_{\mathbb{S},i,j}, W, N^{\mathbb{S}})$ $\rhd$ Eq. 5.8-5.10

8. $\lambda_{\mathbb{S}} \leftarrow \dfrac{1 - E_x^{\mathbb{S}}}{1 - E_x^{\mathbb{S}} + 1 - E_x^{\text{NI}_{\mathcal{S}}} + \sum_{i \in \mathcal{T}} 1 - E^{\text{NI}_i}}$ $\rhd$ Eq. 5.12

9. $r_{\mathbb{S}} \leftarrow \lambda_{\mathbb{S}} \cdot \left(\widetilde{f_{\mathbb{S}}^{W}}\left(Z(x) = Z^{\mathbb{S}'}(x)\big|D_x^{\mathbb{S}'}\right) - \widetilde{f_{\mathbb{S}}^{W}}\left(Z(x) = Z^{\mathbb{S}}(x)\big|D_x^{\mathbb{S}'}\right) \right)$ $\rhd$ Eq. 5.11

10. Reward $\leftarrow$ Reward $+ r_{\mathbb{S}}$

11. If $\text{NI}_{\mathcal{S}}(Neigh_x)$ is available:

  a. $D_x^{\mathcal{S}}, H_x^{\mathcal{S}} \leftarrow Env.\,DataEvent(\text{NI}_{\mathcal{S}}(Neigh_x), N^{\mathcal{S}}, x)$

  b. $N_\zeta^{\mathcal{S}}, \zeta_{\mathcal{S},i,j} \leftarrow Env.\,Replicate(\text{NI}_{\mathcal{S}}(Neigh_x), D_x^{\mathcal{S}}, H_x^{\mathcal{S}})$

  c. $\widetilde{f_{\mathcal{S}}^{W}}(Z(x)\big|D_x^{\mathcal{S}}) \leftarrow Env.\,BuildPDF(N_\zeta^{\mathcal{S}}, \zeta_{\mathcal{S},i,j}, W, N^{\mathcal{S}})$

  d. $\lambda_{\mathcal{S}} \leftarrow \dfrac{1 - E_x^{\text{NI}_{\mathcal{S}}}}{1 - E_x^{\mathbb{S}} + 1 - E_x^{\text{NI}_{\mathcal{S}}} + \sum_{i \in \mathcal{T}} 1 - E^{\text{NI}_i}}$

  e. $r_{\mathcal{S}} \leftarrow \lambda_{\mathcal{S}} \cdot \left(\widetilde{f_{\mathcal{S}}^{W}}\left(Z(x) = Z^{\mathbb{S}'}(x)\big|D_x^{\mathcal{S}}\right) - \widetilde{f_{\mathcal{S}}^{W}}\left(Z(x) = Z^{\mathbb{S}}(x)\big|D_x^{\mathcal{S}}\right) \right)$

12. Reward $\leftarrow$ Reward $+ r_{\mathcal{S}}$

13. For all i in $\mathcal{T}$ do:

  a. If new information $(\text{NI}_i(q_i), q_i)$ is available:

    i. $\lambda_i \leftarrow \dfrac{1 - E^{\text{NI}_i}}{1 - E_x^{\mathbb{S}} + 1 - E_x^{\text{NI}_{\mathcal{S}}} + \sum_{i \in \mathcal{T}} 1 - E^{\text{NI}_i}}$

    ii. $\text{MP}_i^{\mathbb{S}} \leftarrow s_t.\text{AdditionalDataMP}[i]$

    iii. $\text{MP}_i^{\mathbb{S}'} \leftarrow Env.\,ModelPrediction(Track, \mathbb{S}, x, \mathbb{S}^{\mathrm{U}}, \text{NI}_i(q_i), q_i)$

    iv. $\text{MP}.\,append(\text{MP}_i^{\mathbb{S}'})$

    v. $r_i \leftarrow \lambda_i \cdot \left(|\text{MP}_i^{\mathbb{S}} - \text{NI}_i(q_i)| - |\text{MP}_i^{\mathbb{S}'} - \text{NI}_i(q_i)| \right) \cdot \gamma_i^{MP}$ $\rhd$ See Eq. 5.16

    vi. Reward $\leftarrow$ Reward $+ r_i$

Appendix 5.7

This section outlines the algorithm for finding the conditioning data event and its geometry for a mining block.

Input: Number of closest conditioning data $N^\$$, location of mining block x, neighborhood values $Z^\$(Neigh_x)$, and location $Neigh_x$

Data: Tolerance for search $Tol_i^\$$, $\forall i \in N^\$$

Output: Conditioning data event value $D_x^\$$ and its geometry $H_x^\$$

1. $RD \leftarrow Neigh_x - x$ ▷ Finding relative distance vector of neighborhood
2. $L \leftarrow Length(Neigh_x)$ ▷ Length of the neighborhood
3. $B \leftarrow Breadth(Neigh_x)$ ▷ Breadth of the neighborhood
4. $H \leftarrow Height(Neigh_x)$ ▷ Height of the neighborhood
5. $LocN^\$ \leftarrow \min(RD, N^\$)$ ▷ Computing location of $N_D^\$$ points with minimum RD
6. $NC \leftarrow \sqrt{L^2 + B^2 + H^2}$ ▷ Normalizing coefficient for distance vector
7. If $Z^\$(Neigh_x)$ is available:
   a. $D_x^\$ \leftarrow Z^\$(Neigh_x)[LocN^\$]$ ▷ Returning data event value
   b. $H_x^\$ \leftarrow \frac{RD + NC}{2 * NC}$ ▷ Returning normalized distance vector
8. Else:
   a. $D_x^\$ \leftarrow [-1] * N^\$$ ▷ Returning default value
   b. $H_x^\$ \leftarrow [-1] * N^\$$ ▷ Returning default value

---

Appendix 5.8

This section shows the algorithm for generating the state description.

Input: Location of mining block (x), simulation number $(\mathbb{s})$, updated simulation $\mathbb{S}^{U}$

Data: New information $(NI_i(x), i \in \mathcal{S}; q^i, NI_i(q^i), \forall i \in \mathcal{T})$, material flow data $(Track)$, initial stochastic simulations $(\mathbb{S}^{I})$, conditioning simulation data $(N^{\mathbb{s}})$, conditioning spatial sensor data $(N^{\mathcal{S}})$, number of block in neighbourhood in X, Y, and Z direction (nX, nY, nZ), block size in X, Y, and Z direction (sX, sY, sZ)

Output: State (s_t) and neighborhood block location $(Neigh_x)$

1. $Neigh_x \leftarrow []$
2. For i in $(-nX, nX, sX)$ do
  a. For i in $(-nY, nY, sY)$ do
    1. For i in $(-nZ, nZ, sZ)$ do
      i. $Neigh_x.append([i + x[0], j + x[1], k + x[2]])$
3. SimState $\leftarrow Env.reshape(Z^{\mathbb{s}}(Neigh_x)), \forall \mathbb{s} \in \mathbb{S}^{I}$ ▷ Image input
4. If spatial new information $(NI_{\mathcal{S}})$ is available:
  a. SpatialState $\leftarrow Env.reshape(NI_{\mathcal{S}}(Neigh_x))$ ▷ Image input
  b. AdditionalDataSpatial $\leftarrow Env.DataEvent(NI_{\mathcal{S}}(Neigh_x), N^{\mathcal{S}}, x, Neigh_x)$
5. AdditionalDataSim $\leftarrow Env.DataEvent(Z^{\mathbb{s}'}(Neigh_x), N^{\mathbb{s}}, x, Neigh_x), \forall \mathbb{s}' \in \mathbb{S}^{U}$
6. AdditionalDataES $\leftarrow E_x^{\mathbb{s}}$ ▷ Eq. 5.1
7. AdditionalDataEB $\leftarrow E_x^{NI_{\mathcal{S}}}$ ▷ Eq. 5.2
8. If temporal new information $(q_i, NI_i(q_i))$ is available:
  a. AdditionalDataEI $\leftarrow []$, AdditionalDataOS $\leftarrow []$, AdditionalDataMP $\leftarrow []$
    a. For i in $\mathcal{T}$ do:
      1. AdditionalDataEI. $\leftarrow (E^{NI_i})$
      2. AdditionalDataOS. $\leftarrow (NI_i(q_i), q_i)$
      3. AdditionalDataMP. $\leftarrow$
        $(Env.ModelPrediction(Track, \mathbb{s}, x, \mathbb{S}^{I}, NI_i(q_i), q_i))$
  b. Else
    a. For i in $\mathcal{T}$ do:
      1. AdditionalDataEI. $\leftarrow (-1)$
      2. AdditionalDataOS. $\leftarrow (-1, -1)$
      3. AdditionalDataMP. $\leftarrow (-1)$
9. $s_t \leftarrow$ (SimState, SpatialState, AdditionalDataSim, AdditionalDataSpatial , AdditionalDataES, AdditionalDataEB , AdditionalDataEI , AdditionalDataOS , AdditionalDataMP)

Appendix 5.9

 This section details the architecture of the agents used for the case study in the manuscript. The total number of parameters for the actor agent is ≈ 497,000, and the critic agent is ≈ 498,000.

 The actor agent consists of:

1. A simulation state and a spatial sensor data state, each having the following convolution layers:
    a. Convolution layer 1: 8 filters, kernel size 5, strides 1, ReLU activation, valid padding and Xavier initialization (Glorot and Bengio, 2010)
    b. Convolution layer 2: 8 filters, kernel size 3, strides 1, ReLU activation, valid padding and Xavier initialization
    c. Dense layer 1: Takes flattened convolution output as input and have 128 neurons with ReLU activation and Xavier initialization

 The outputs from the convolution layers are concatenated with the additional data. The actor agent then has the following layers:

2. Dense layer 2: 100 neurons with Xavier initialization and ReLU activation
3. Dense layer 3: 50 neurons with Xavier initialization and ReLU activation
4. Dense layer 4: 1 output with Xavier initialization and sigmoid activation

 The critic agent consisted of:

1. A simulation state and a spatial sensor data state, each having the following convolution layers:
    a. Convolution layer 1: 8 filters, kernel size 5, strides 1, ReLU activation, valid padding and Xavier initialization (Glorot and Bengio, 2010)
    b. Convolution layer 2: 8 filters, kernel size 3, strides 1, ReLU activation, valid padding and Xavier initialization

    c. Dense layer 1: Takes flattened convolution output as input and have 128 neurons with ReLU activation and Xavier initialization

The outputs from the convolution layers are concatenated with the additional data and the action. The critic agent then has the following layers:

2. Dense layer 2: 100 neurons with Xavier initialization and ReLU activation
3. Dense layer 3: 50 neurons with Xavier initialization and ReLU activation
4. Dense layer 4: 1 output with Xavier initialization and no activation

Appendix 5.10

The proposed AI algorithm is trained on an Intel® i7-8700 machine with an 8-core processor and an NVIDIA GeForce GTX 1050 GPU for approximately 2 days. The training process was executed with 100 training episodes (N_{TE}), a training interval of 65000 (N_I), training iterations of 200 (N_T), a batch size of 500 (N_{BS}), 20 initial geostatistically simulated models ($\mathbb{S}^I$), an L2 regularization cost of 0.001 (c), a discount factor of 0.99 (γ), a soft update parameter of 0.001 (τ), an adjustment factor for initial geostatistical simulation conditional variance of 10 ($\gamma_{\mathbb{S}}$), a model-based prediction error adjustment factor of 10 ($\gamma_i^{\mathrm{MP}}, \forall i \in \mathcal{T}$), a Legendre polynomial order of 10 (W), a number of conditioning data events in the geostatistically simulated model of 8 ($N^{\mathbb{S}}$), and a replay memory buffer size of 1 million (N_R). The noise process used for the exploration during the training phase of the algorithm is the Ornstein-Uhlenbeck method, with a standard deviation of 0.1 (Uhlenbeck and Ornstein, 1930). The learning rates of the actor and critic are set to 0.0001 and 0.001, respectively. Most of the parameters used in the proposed algorithm are taken from Lillicrap et al. (2015).

CHAPTER 6
Conclusions and Future Work

6.1   Conclusions

A mining complex is an integrated value chain where raw materials flow from mines to market after being treated and transformed into sellable products through the available processing streams. The simultaneous stochastic optimization model for long-term production planning of a mining complex aims to jointly optimize the different strategic decisions in a single optimization model to maximize asset valuation while accounting for uncertainty in the supply of materials and commodity prices. The single optimization model capitalizes on synergies between the different components and the quantified variability and uncertainty, to better meet the production targets. The existing simultaneous stochastic optimization models do not account for supply uncertainty related to the geometallurgical properties of materials. Geometallurgical properties of materials such as hardness, grindability, ore texture, and so on, affect the throughput and energy consumption of the processing streams in a mining complex. A short-term production plan is then generated within the predefined long-term production plan to provide monthly/weekly/daily time scale production decisions that aim to ensure compliance with the long-term targets while maximizing cash flow. Conventional and new digital technologies, including the development of advanced sensors, monitoring, and tracking devices, have enabled a mining complex to acquire new information about the performance of its different

248

components, and about pertinent properties of materials at those components. However, the new information is "soft," i.e., noisy, and therefore uncertain, because of the characteristics of the related sensors that generate indirect measurements compared to those derived from the analysis of the drillhole samples in geochemical laboratories. The existing technologies cannot integrate the noisy and uncertain incoming new information and respond accordingly to adapt the short-term production plan of a mining complex.

This thesis presents a new self-learning artificial intelligence framework to allow a mining complex to learn to integrate noisy and uncertain incoming new information and respond quickly by adapting the short-term production planning decisions. This follows an initial general effort to extend the simultaneous stochastic optimization model to account for the geometallurgical properties of materials.

Chapter 2 presents an approach to include geometallurgical supply uncertainty and decisions in the simultaneous stochastic optimization of the long-term production plan of a mining complex. This approach characterizes the materials as hard and soft, depending on their geometallurgical properties, and then links those properties to the simultaneous stochastic optimization model through geometallurgical targets. The contribution and applied aspects of the proposed approach are highlighted with an application at a large copper-gold mining complex. Two geometallurgical properties related to the grindability of the material, i.e., the semi-autogenous power index and the bond work index, are considered in the optimization model. The forecasts for the stochastic mine production plan show a 12.5% higher production of the primary copper product, 22.9%, 32.4%, and 34.7% higher production of gold, silver, and molybdenum secondary products, higher chances of meeting the different production

targets, and a 19.3% increase in net present value compared to the conventional mine production plan.

Chapter 3 presents a new self-learning artificial intelligence framework for adapting the short-term flow of materials in a mining complex with incoming new information. The framework consists of two parts: first, updating the supply uncertainty related to multiple correlated properties of materials with a new extension of ensemble Kalman filter and, second, training neural agents to respond to incoming new information by adapting the short-term flow of materials in multiple product mining complexes with a model-free policy gradient reinforcement learning algorithm. The short-term flow of materials includes deciding which destinations to send the extracted materials and then determining how to utilize the processing streams. The proposed framework is applied at a real-world copper mining complex to show its applicability in an operating mine, and its ability to adapt the flow of materials in real-time to more closely meet the production requirements while increasing the production of primary copper products by 11% and cash flow by 15% compared to industry-standard approaches.

The major short-term production decisions in a mining complex, such as extraction sequence, destination policies, and processing stream utilization, are interrelated and dependent on the performance of its different components. Therefore, in Chapter 4, a new model-based self-play reinforcement learning algorithm is proposed to adapt all the major short-term production planning decisions simultaneously by responding to the incoming new information. The proposed algorithm plays the game of short-term production planning by itself using a Monte Carlo tree search, to train a deep neural network agent to adapt the short-term

250

production plan by responding to incoming new information in an operating mine. An application of the proposed algorithm at a real-world copper mining complex shows its applicability in an operating mine and its ability to adapt the short-term production plan in real-time to more closely meet the production requirements while increasing copper production by 7% and cash flow by 12%.

The existing methods to update the supply uncertainty of materials do not learn from the incoming new information, do not capitalize on the spatial statistics in the incoming spatial new information (collected spatially with drilling machines located in the mineral deposits), and do not account for nor respect high-order spatial statistics. Therefore, in Chapter 5, a new model-free actor-critic reinforcement learning algorithm is proposed that leverages concepts of high-order spatial statistics to train convolutional neural network agents to update the supply uncertainty of materials by integrating the incoming new information. An application of the proposed algorithm at a copper mining operation with incoming drilling machine sensor data (spatial), and processing mill sensor data (temporal) demonstrates its applied aspects in updating the supply uncertainty of materials in real-time while reproducing spatial patterns and high-order spatial statistics.

This proposed self-learning artificial intelligence framework will allow a mining complex to learn from and adapt to incoming new information while it operates, to make more informed short-term production planning decisions and more closely meet its long-term production plan and targets.

6.2   Future Work

There are several possible directions for future research based on the work presented in this thesis. The approach outlined in Chapter 2 to integrate geometallurgical properties of materials in the simultaneous stochastic optimization model for long-term planning could be further developed to integrate more geometallurgical properties of materials and operating mode related decisions. The results from the case study in Chapter 2 indicate that the available destinations and processing streams may be better utilized, providing further opportunities for value improvement through the inclusion of capital expenditure decisions in the simultaneous stochastic optimization model.

The self-play reinforcement learning algorithm proposed in Chapter 4 could be improved to match the performance of recent algorithms proposed in Silver et al. (2017, 2018) and Schrittwieser et al. (2019). It could also be tested at different mining complexes with different sources of incoming new information to evaluate its robustness along with a sensitivity analysis of the parameters used in proposed algorithm. Future research could also investigate incorporating supply uncertainty and incoming new information related to the geometallurgical properties of materials in the proposed algorithm. The algorithm could be extended in the future to adapt the fleet allocation and assignment decisions in addition to the major short-term planning decisions it already considers. In the future, this algorithm could also be integrated into the hyper-heuristic algorithms used for solving simultaneous stochastic optimization models of mining complexes. The adaptation of the short-term production plan within a predefined long-term production plan limits the scope for significant improvements and adaptation; therefore, future research could also focus on adapting the short and long-term production plans simultaneously by responding to incoming new information. Mining complexes consist of different components that

compete for resources and cooperate to generate products. A multi-agent architecture could be explored in the future to assign different agents to different short-term production planning decisions that compete and cooperate, like the algorithm proposed in Baker et al. (2019).

The actor-critic reinforcement learning algorithm proposed in Chapter 5 could be improved to match the performance of advantage, and asynchronous advantage actor-critic algorithms proposed in Mnih et al. (2016). In the future, research could focus on integrating additional sources of incoming new information, applying this algorithm at different mining operations, sensitivity analysis of the parameters, and expanding it to update multiple properties of materials at block support scale. In the future, the proposed algorithm could also be expanded and applied to reservoirs and oil fields. It could also be extended to update the categorical geological, and non-additive geometallurgical properties of materials.

References

Aanonsen S, Nævdal G, Oliver D, Reynolds A, Vallès B (2009) The ensemble Kalman Filter in reservoir engineering: A Review. *SPE Journal* 14:393–412

Abzalov MZ, 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* 119:132–142. doi: 10.1179/1743275811Y.0000000007

Aissani N, Bekrar A, Trentesaux D, Beldjilali B (2012) Dynamic scheduling for multi-site companies: A decisional approach based on reinforcement multi-agent learning. *Journal of Intelligent Manufacturing* 23:2513–2529. doi: 10.1007/s10845-011-0580-y

Alarie S, Gamache M (2002) Overview of solution strategies used in truck dispatching systems for open pit mines. *International Journal of Surface Mining, Reclamation and Environment* 16:59–76. doi: 10.1076/ijsm.16.1.59.3408

Albor Consuegra FR, Dimitrakopoulos R (2009) Stochastic mine design optimisation based on simulated annealing: pit limits, production schedules, multiple orebody scenarios and sensitivity analysis. *Mining Technology* 118:79–90. doi: 10.1179/037178409X12541250836860

Alruiz OM, Morrell S, Suazo CJ, Naranjo A (2009) A novel approach to the geometallurgical modelling of the Collahuasi grinding circuit. *Minerals Engineering* 22:1060–1067. doi: 10.1016/j.mineng.2009.03.017

Amini H, Gholami R, Monjezi M, Torabi SR, Zadhesh J (2012) Evaluation of flyrock phenomenon due to blasting operation by support vector machine. *Neural Computing and Applications* 21:2077–2085. doi: 10.1007/s00521-011-0631-5

Arel I, Liu C, Urbanik T, Kohls AG (2010) Reinforcement learning-based multi-agent system for network traffic signal control. *IET Intelligent Transport Systems* 4:128–135. doi: 10.1049/iet-its.2009.0070

Arpat GB, Caers J (2007) Conditional simulation with patterns. *Mathematical Geology* 39:177–203. doi: 10.1007/s11004-006-9075-3

Arulkumaran K, Deisenroth MP, Brundage M, Bharath AA (2017) Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine* 34:26–38. doi: 10.1109/MSP.2017.2743240

Asad MWA, Dimitrakopoulos R (2013) A heuristic approach to stochastic cutoff grade optimization for open pit mining complexes with multiple processing streams. *Resources Policy* 38:591–597. doi: 10.1016/j.resourpol.2013.09.008

Asad MWA, Qureshi MA, Jang H (2016) A review of cut-off grade policy models for open pit mining operations. *Resources Policy* 49:142–152. doi: 10.1016/j.resourpol.2016.05.005

Avalos S, Ortiz JM (2019) Geological modeling using recursive convolutional neural networks. *arXiv preprint* 1–8. doi: https://arxiv.org/abs/1904.12190

Babak O, Deutsch C V. (2009) Collocated cokriging based on merged secondary attributes. *Mathematical Geosciences* 41:921–926. doi: 10.1007/s11004-008-9192-2

Baek J, Choi Y (2019) Simulation of truck haulage operations in an underground mine using big data from an ICT-based mine safety management system. *Applied Sciences* 9:2639. doi: 10.3390/app9132639

Bahrami S, Doulati Ardejani F, Baafi E (2016) Application of artificial neural network coupled with genetic algorithm and simulated annealing to solve groundwater inflow problem to an advancing open pit mine. *Journal of Hydrology* 536:471–484. doi: 10.1016/j.jhydrol.2016.03.002

Baker B, Kanitscheider I, Markov T, Wu Y, Powell G, McGrew B, Mordatch I (2019) Emergent tool use from multi-agent autocurricula. *arXiv preprint* 1–28. doi: http://arxiv.org/abs/1909.07528

Ballantyne GR, Powell MS (2014) Benchmarking comminution energy consumption for the processing of copper and gold ores. *Minerals Engineering* 65:109–114. doi: 10.1016/j.mineng.2014.05.017

Bandarian EM, Bloom LM, Mueller UA (2008) Direct minimum/maximum autocorrelation factors within the framework of a two structure linear model of coregionalisation. *Computers and Geosciences* 34:190–200. doi: 10.1016/j.cageo.2007.03.015

Barbaro R, Ramani R (1986) Generalized multiperiod MIP model for production scheduling and processing facilities selection and location. *Mining Engineering* 38:107–114

Barde SRA, Yacout S, Shin H (2019) Optimal preventive maintenance policy based

on reinforcement learning of a fleet of military trucks. *Journal of Intelligent Manufacturing* 30:147–161. doi: 10.1007/s10845-016-1237-7

Barnett RM, Deutsch C V. (2012) Practical implementation of non-linear transforms for modeling geometallurgical variables. In: Geostatistics Oslo. pp 409–422

Benndorf J (2015) Making use of online production data: Sequential updating of mineral resource models. *Mathematical Geosciences* 47:547–563. doi: 10.1007/s11004-014-9561-y

Benndorf J (2020) Data Assimilation for Resource Model Updating. In: SpringerBriefs in Applied Sciences and Technology. Springer, pp 19–60

Benndorf J, Buxton MWN (2016) Sensor-based real-time resource model reconciliation for improved mine production control: A conceptual framework. *Mining Technology* 125:54–64. doi: 10.1080/14749009.2015.1107342

Benndorf J, Dimitrakopoulos R (2013) Stochastic long-term production scheduling of iron ore deposits: Integrating joint multi-element geological uncertainty. *Journal of Mining Science* 49:68–81. doi: 10.1134/S1062739149010097

Bette H, Noll R, Müller G, Jansen H-W, Nazikkol Ç, Mittelstädt H (2005) High-speed scanning laser-induced breakdown spectroscopy at 1000 Hz with single pulse evaluation for the detection of inclusions in steel. *Journal of Laser Applications* 17:183–190. doi: 10.2351/1.1961738

Bhuiyan M, Esmaieli K, Ordóñez-Calderón JC (2019) Application of data analytics techniques to establish geometallurgical relationships to bond work index at the Paracutu mine, Minas Gerais, Brazil. *Minerals* 9:302. doi: 10.3390/min9050302

Bielli M, Bielli A, Rossi R (2011) Trends in models and algorithms for fleet management. *Procedia - Social and Behavioral Sciences* 20:4–18. doi: 10.1016/j.sbspro.2011.08.004

Blom M, Pearce AR, Stuckey PJ (2018) Multi-objective short-term production scheduling for open-pit mines: a hierarchical decomposition-based algorithm. *Engineering Optimization* 50:2143–2160. doi: 10.1080/0305215X.2018.1429601

Blom M, Pearce AR, Stuckey PJ (2017) Short-term scheduling of an open-pit mine with multiple objectives. *Engineering Optimization* 49:777–795. doi: 10.1080/0305215X.2016.1218002

Blom M, Pearce AR, Stuckey PJ (2019) Short-term planning for open pit mines: A review. *International Journal of Mining, Reclamation and Environment* 33:318–339. doi: 10.1080/17480930.2018.1448248

Boisvert JB, Rossi ME, Ehrig K, Deutsch C V (2013) Geometallurgical modeling at Olympic Dam mine, South Australia. *Mathematical Geosciences* 45:901–925. doi: 10.1007/s11004-013-9462-5

Boland N, Dumitrescu I, Froyland G (2008) A multistage stochastic programming approach to open pit mine production scheduling with uncertain geology. *Optimization Online* 1–33

Bolger JA (2000) Semi-quantitative laser-induced breakdown spectroscopy for analysis of mineral drill core. *Applied Spectroscopy* 54:181–189. doi: 10.1366/0003702001949375

Bond FC (1952) Third theory of comminution. *Mining Enginnering* 4:484–494

Both C, Dimitrakopoulos R (2020) Joint stochastic short-term production scheduling and fleet management optimization for mining complexes. *Optimization and Engineering* 1–27. doi: https://doi.org/10.1007/s11081-020-09495-x

Bottou L (2010) Large-scale machine learning with stochastic gradient descent. In: Proceedings of COMPSTAT'2010. pp 177–186

Boucher A, Dimitrakopoulos R (2009) Block simulation of multiple correlated variables. *Mathematical Geosciences* 41:215–237. doi: 10.1007/s11004-008-9178-0

Boucher A, Dimitrakopoulos R (2012) Multivariate block-support simulation of the Yandi iron ore deposit, Western Australia. *Mathematical Geosciences* 44:449–468. doi: 10.1007/s11004-012-9402-9

Boulter A, Hall R (2015) Wireless network requirements for the successful implementation of automation and other innovative technologies in open-pit mining. *International Journal of Mining, Reclamation and Environment* 29:368–379. doi: 10.1080/17480930.2015.1086549

Brewer A, Nancy S, Thomas L (1999) Intelligent tracking in manufacturing. *Journal of Intelligent Manufacturing* 10:245–250. doi: 10.1023/A:1008995707211

Brown GR, Hwang PYC (1992) Introduction to random signals and applied Kalman filtering, Volume 3. New York: Wiley

Bu X, Rao J, Xu CZ (2009) A reinforcement learning approach to online web systems auto-configuration. In: International Conference on Distributed Computing Systems. pp 2–11

Campeau LP, Dubois M (2019) Artificial intelligence using real-time data. In: Application of Computers and Operations Research in the Mineral Industry. pp 557–562

Chaowasakoo P, Leelasukseree C, Wongsurawat W (2014) Introducing GPS in fleet management of a mine: Impact on hauling cycle time and hauling capacity. *International Journal of Technology Intelligence and Planning* 10:49–66. doi: 10.1504/IJTIP.2014.066711

Chaowasakoo P, Seppälä H, Koivo H, Zhou Q (2017) Digitalization of mine operations: Scenarios to benefit in real-time truck dispatching. *International Journal of Mining Science and Technology* 27:229–236. doi: 10.1016/j.ijmst.2017.01.007

Chatterjee S (2013) Vision-based rock-type classification of limestone using multi-class. *Applied intelligence* 39:14–27. doi: 10.1007/s10489-012-0391-7

Chatterjee S, Dash A, Bandopadhyay S (2015) Ensemble support vector machine algorithm for reliability estimation of a mining machine. *Quality and Reliability Engineering International* 31:1503–1516. doi: 10.1002/qre.1686

Chatterjee S, Dimitrakopoulos R (2012) Multi-scale stochastic simulation with a wavelet-based approach. *Computers and Geosciences* 45:177–189. doi: 10.1016/j.cageo.2011.11.006

Chatterjee S, Dimitrakopoulos R, Mustapha H (2012) Dimensional reduction of pattern-based simulation using wavelet analysis. *Mathematical Geosciences* 44:343–374. doi: 10.1007/s11004-012-9387-4

Chatterjee S, Mustapha H, Dimitrakopoulos R (2016) Fast wavelet-based stochastic simulation using training images. *Computational Geosciences* 20:399–420. doi: 10.1007/s10596-015-9482-y

Chen Y, Oliver DS (2012) Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Mathematical Geosciences* 44:1–26. doi: 10.1007/s11004-011-9376-z

Coward SJV, Dunham S, Stewart M (2009) The primary-response framework for

geometallurgical variables. In: Seventh international mining geology conference. pp 109–113

Cremers DA (1987) Analysis of metals at a distance using laser-induced breakdown spectroscopy. *Applied Spectroscopy* 41:572–579. doi: 10.1366/0003702874448742

Curry JA, Ismay MJL, Jameson GJ (2014) Mine operating costs and the potential impacts of energy and grinding. *Minerals Engineering* 56:70–80. doi: 10.1016/j.mineng.2013.10.020

Dalm M, Buxton MWN, van Ruitenbeek FJA (2018) Ore–waste discrimination in epithermal deposits using near-infrared to short-wavelength infrared (NIR-SWIR) hyperspectral imagery. *Mathematical Geosciences* 51:1–27. doi: 10.1007/s11004-018-9758-6

Dalm M, Buxton MWN, van Ruitenbeek FJA (2017) Discriminating ore and waste in a porphyry copper deposit using short-wavelength infrared (SWIR) hyperspectral imagery. *Minerals Engineering* 105:10–18. doi: 10.1016/j.mineng.2016.12.013

Dalm M, Buxton MWN, Van Ruitenbeek FJA, Voncken JHL (2014) Application of near-infrared spectroscopy to sensor based sorting of a porphyry copper ore. *Minerals Engineering* 58:7–16. doi: 10.1016/j.mineng.2013.12.016

Dassault S (2017) Geovia Whittle, strategic mine planning software

David M (1988) Handbook of applied advanced geostatistical ore reserve estimation. Elsevier, Amsterdam

Davis BM, Greenes KA (1983) Estimation using spatially distributed multivariate data: An example with coal quality. *Journal of the International Association for Mathematical Geology* 15:287–300. doi: 10.1007/BF01036071

Davis MW (1987) Production of conditional simulations via the LU triangular decomposition of the covariance matrix. *Mathematical Geology* 19:91–98. doi: 10.1007/BF00898189

de Carvalho J, Dimitrakopoulos R (2019) Effects of high-order simulations on the simultaneous stochastic optimization of mining complexes. *Minerals* 9:210. doi: 10.3390/min9040210

de Carvalho JP, Dimitrakopoulos R, Minniakhmetov I (2019) High-order block

support spatial simulation method and its application at a gold deposit. *Mathematical Geosciences* 51:793–810. doi: 10.1007/s11004-019-09784-x

De Jong TPR (2004) Automatic sorting of minerals. In: IFAC Proceedings. pp 441–446

Death DL, Cunningham AP, Pollard LJ (2009) Multi-element and mineralogical analysis of mineral ores using laser induced breakdown spectroscopy and chemometric analysis. *Spectrochimica Acta - Part B* 64:1048–1058. doi: 10.1016/j.sab.2009.07.017

Death DL, Cunningham AP, Pollard LJ (2008) Multi-element analysis of iron ore pellets by laser-induced breakdown spectroscopy and principal components regression. *Spectrochimica Acta - Part B Atomic Spectroscopy* 63:763–769. doi: 10.1016/j.sab.2008.04.014

Del Castillo MF (2018) Dynamic simultaneous optimization of mineral value chains under resource uncertainty. Thesis, McGill University.

Del Castillo MF, Dimitrakopoulos R (2019) Dynamically optimizing the strategic plan of mining complexes under supply uncertainty. *Resources Policy* 60:83–93. doi: 10.1016/j.resourpol.2018.11.019

Del Castillo MF, Dimitrakopoulos R (2016) A multivariate destination policy for geometallurgical variables in mineral value chains using coalition-formation clustering. *Resources Policy* 50:322–332. doi: 10.1016/j.resourpol.2016.10.003

Demirel N, Emil MK, Duzgun HS (2011) Surface coal mine area monitoring using multi-temporal high-resolution satellite imagery. *International Journal of Coal Geology* 86:3–11. doi: 10.1016/j.coal.2010.11.010

Desbarats AJ, Dimitrakopoulos R (2000) Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors. *Mathematical Geology* 32:919–942. doi: 10.1023/A:1007570402430

Deutsch C, Journel AG (1992) GSLIB: Geostatistical software library and user's guide, 2nd edn. Oxford University Press, New York

Deutsch JL, Palmer K, Deutsch C V., Szymanski J, Etsell TH (2016) Spatial modeling of geometallurgical properties: Techniques and a case study. *Natural Resources Research* 25:161–181. doi: 10.1007/s11053-015-9276-x

Dimitrakopoulos R (2011) Stochastic optimization for strategic mine planning: A

decade of developments. *Journal of Mining Science* 47:138–150. doi: 10.1134/S1062739147020018

Dimitrakopoulos R, Farrelly CT, Godoy M (2002) Moving forward from traditional optimisation: grade uncertainty and risk effects in open pit design. *Mining Technology* 111:82–88. doi: 10.1179/mnt.2002.111.1.82

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

Dimitrakopoulos R, Kaklis K (2001) Integration of assay and cross-hole tomographic data in orebody modelling: Joint geostatistical simulation and application at Mount Isa mine, Queensland. *Applied Earth Science* 110:33–39. doi: 10.1179/aes.2001.110.1.33

Dimitrakopoulos R, Luo X (2004) Generalized sequential Gaussian simulation on group size v and screen-effect approximations for large field simulations. *Mathematical Geology* 36:567–590. doi: 10.1023/B:MATG.0000037737.11615.df

Dimitrakopoulos R, Mustapha H, Gloaguen E (2010) High-order statistics of spatial random fields: Exploring spatial cumulants for modeling complex non-Gaussian and non-linear phenomena. *Mathematical Geosciences* 42:65–99. doi: 10.1007/s11004-009-9258-9

Dimitrakopoulos R, Ramazan S (2005) Uncertainty-based production scheduling in open pit mining. *SME Transactions* 316:106–112

Dovera L, Della Rossa E (2011) Multimodal ensemble Kalman filtering using Gaussian mixture models. *Computational Geosciences* 15:307–323. doi: 10.1007/s10596-010-9205-3

Dowd P (1994) Risk assessment in reserve estimation and open-pit planning. *Mining Technology* 103:148–154. doi: 10.1016/0148-9062(95)97056-O

Dowd PA (1997) Risk in minerals projects: analysis, perception and management. *Mining Technology* 106:9–18

Dowd PA, Xu C, Coward S (2016) Strategic mine planning and design: Some challenges and strategies for addressing them. *Mining Technology* 125:22–34.

doi: 10.1179/1743286315Y.0000000032

Dunham SJV, Coward S (2011) Beyond geometallurgy—gaining competitive advantage by exploiting the broad view of geometallurgy. In: International Geometallurgy Conference. pp 5–7

Eilbert RF (2009) X-ray Technologies. In: Aspects of Explosives Detection. Elsevier, pp 89–130

Eivazy H, Askari-Nasab H (2012) A mixed integer linear programming model for short-term open pit mine production scheduling. *Mining Technology* 121:97–108. doi: 10.1179/1743286312Y.0000000006

Ercelebi SG, Bascetin A (2009) Optimization of shovel-truck system for surface mining. *Journal of the Southern African Institute of Mining and Metallurgy* 109:433–439

Evensen G (1994a) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research* 99:10143–10162. doi: 10.1029/94jc00572

Evensen G (2009) Data assimilation: The ensemble Kalman filter. Springe, Berlin, Heidelberg, 9783642037108

Evensen G (1994b) Inverse methods and data assimilation in nonlinear ocean models. *Physica D: Nonlinear Phenomena* 77:108–129. doi: 10.1016/0167-2789(94)90130-9

Farmer IW (2016) Stochastic mining supply chain optimization: A study of integrated capacity decisions and pushback design under uncertainty. Thesis, McGill University.

Farzanegan A, Vahidipour SM (2009) Optimization of comminution circuit simulations based on genetic algorithms search method. *Minerals Engineering* 22:719–726. doi: 10.1016/j.mineng.2009.02.009

Fichtl P, Fournier F, Royer JJ (1997) Cosimulations of lithofacies and associated reservoir properties using well and seismic data. In: SPE Annual Technical Conference and Exhibition. pp 381–393

Franco-Sepúlveda G, Del Rio-Cuervo JC, Pachón-Hernández MA (2019) State of the art about metaheuristics and artificial neural networks applied to open pit mining. *Resources Policy* 60:125–133. doi: 10.1016/j.resourpol.2018.12.013

Fu J, Gómez-Hernández JJ, Du S (2017) A gradient-based blocking Markov Chain Monte Carlo method for stochastic inverse modeling. In: Quantitative Geology and Geostatistics. Springer, Cham, pp 777–788

Fytas K, Calder PN (1986) A computerized model of open pit short and long range production scheduling. In: Application of Computers and Operations Research in the Mineral Industry. pp 109–119

Fytas K, Hadjigeorgiou J, Collins JL (1993) Production scheduling optimization in open pit mines. *International Journal of Surface Mining, Reclamation and Environment* 7:1–9. doi: 10.1080/09208119308964677

Ganin Y, Kulkarni T, Babuschkin I, Eslami SMA, Vinyals O (2019) Synthesizing programs for images using reinforced adversarial learning. *arXiv preprint* 1–12. doi: https://arxiv.org/abs/1804.01118

Garrido M, Ortiz JM, Villaseca F, Kracht W, Townley B, Miranda R (2019) Change of support using non-additive variables with Gibbs Sampler: Application to metallurgical recovery of sulphide ores. *Computers and Geosciences* 122:68–76. doi: 10.1016/j.cageo.2018.10.002

Gemcom (2005) Surface mining and design. In: Whittle Strategic Mine Planning

Gershon ME (1983) Mine scheduling optimization with mixed integer programming. *Mining Engineering* 35:351–354

Gholamnejad J (2008) A zero-one integer programming model for open pit mining sequences. *Journal of the Southern African Institute of Mining and Metallurgy* 108:759–762

Gilani SO, Sattarvand J (2016) Integrating geological uncertainty in long-term open pit mine production planning by ant colony optimization. *Computers and Geosciences* 87:31–40. doi: 10.1016/j.cageo.2015.11.008

Gilman J, Ozgen C (2013) Reservoir simulation: history matching and forecasting. Society of Petroleum Engineers, Texas

Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the 13th international conference on artificial intelligence and statistics. pp 249–256

Godoy M (2002) The effective management of geological risk in long-term production scheduling of open pit mines. Thesis, The University of Queensland.

Godoy M, Dimitrakopoulos R (2004) Managing risk and waste mining in long-term production scheduling of open-pit mines. In: SME Transactions. pp 43–50

Goetz AFH, Curtiss B, Shiley DA (2009) Rapid gangue mineral concentration measurement over conveyors by NIR reflectance spectroscopy. *Minerals Engineering* 22:490–499. doi: 10.1016/j.mineng.2008.12.013

Goodfellow R (2014) Unified modeling and simultaneous optimization of open pit mining complexes with supply uncertainty. Thesis, McGill University.

Goodfellow R, Albor Consuegra F, Dimitrakopoulos R, Lloyd T (2012) Quantifying multi-element and volumetric uncertainty, Coleman McCreedy deposit, Ontario, Canada. *Computers and Geosciences* 42:71–78. doi: 10.1016/j.cageo.2012.02.018

Goodfellow R, Dimitrakopoulos R (2015) Stochastic optimization of open pit mining complexes with capital expenditures: Application at a copper mining complex. In: Application of Computers and Operations Research in the Mineral Industry. pp 657–667

Goodfellow R, Dimitrakopoulos R (2016) Global optimization of open pit mining complexes with uncertainty. *Applied Soft Computing* 40:292–304. doi: 10.1016/j.asoc.2015.11.038

Goodfellow R, Dimitrakopoulos R (2017) Simultaneous stochastic optimization of mining complexes and mineral value chains. *Mathematical Geosciences* 49:341–360. doi: 10.1007/s11004-017-9680-3

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press Inc. New York

Grønnevik R, Evensen G (2001) Application of ensemble-based techniques in fish stock assessment. *Sarsia* 86:517–526. doi: 10.1080/00364827.2001.10420490

Guardiano FB, Srivastava RM (1993) Multivariate geostatistics: Beyond bivariate moments. In: Geostatistics Troia '92 Quantitative Geology and Geostatistics. Kluwer; Quantitative Geology and Geostatistics, 5, pp 133–144

Guo X, Lee H, Wang X, Lewis R (2014) Deep learning for real-time Atari game play using offline Monte-Carlo tree search planning. In: Advances in neural information processing systems. pp 3338–3346

Gutiérrez-Esparza JC, Gómez-Hernández JJ (2017) Inverse modeling aided by the

classification and regression tree (CART) algorithm. In: Quantitative Geology and Geostatistics. Springer, Cham, pp 805–819

Hastie T, Tibshirani R, Jerome F (2009) The elements of statistical learning, Springer series in statistics, 2nd edn. Springer

Hausknecht M, Stone P (2015) Deep recurrent Q-learning for partially observable MDPs. In: Association for the Advancement of Artificial Intelligence Fall Symposium. pp 29–37

He N, Reynolds AC, Oliver DS (1997) Three-dimensional reservoir description from multiwell pressure data and prior information. *SPE Journal* 2:312–327

Heidari L, Gervais V, Ravalec M Le, Wackernagel H (2013) History matching of petroleum reservoir models by the ensemble Kalman filter and parameterization methods. *Computers and Geosciences* 55:84–95. doi: 10.1016/j.cageo.2012.06.006

Hinton GE, Srivastava N, Swersky K (2012) Neural netwrok for machine learning-Lecture 6a: Overview of mini-batch gradient descent

Hoerger S, Hoffman L, Seymour F (1999) Mine planning at Newmont's Nevada operations. *Mining engineering* 51:26–30

Horta A, Soares A (2010) Direct sequential Co-simulation with joint probability distributions. *Mathematical Geosciences* 42:269–292. doi: 10.1007/s11004-010-9265-x

Hotelling H (1933) Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24:417–441. doi: 10.1037/h0071325

Hou J, Zhou K, Zhang XS, Kang XD, Xie H (2015) A review of closed-loop reservoir management. *Petroleum Science* 12:114–128. doi: 10.1007/s12182-014-0005-6

Houtekamer P., Mitchell HL (2001) A sequential ensemble Kalman filter for atmospheric data assimilation. *American Meteorological Society* 129:123–137

Hu LY (2000) Gradual deformation and iterative calibration of Gaussian-related stochastic models. *Mathematical Geology* 32:87–108. doi: 10.1023/A:1007506918588

Hu LY, Zhao Y, Liu Y, Scheepens C, Bouchard A (2013) Updating multipoint

simulations using the ensemble Kalman filter. *Computers and Geosciences* 51:7–15. doi: 10.1016/j.cageo.2012.08.020

Huang Z, Cai W, Banfield AF (2009) A new short-and medium-term production planning tool–Minesight® Schedule optimizer. In: SME Transactions. pp 1–5

Huang Z, Heng W, Zhou S (2019) Learning to paint With model-based deep reinforcement learning. In: Proceedings of the IEEE International Conference on Computer Vision. pp 8709–8718

Hustrulid W, Kuchta M, Martin R (2013) Open pit mine planning and design. CRC Press

Iphar M, Yavuz M, Ak H (2008) Prediction of ground vibrations resulting from the blasting operations in an open-pit mine by adaptive neuro-fuzzy inference system. *Environmental Geology* 56:97–107. doi: 10.1007/s00254-007-1143-6

Isaaks EH, Srivastava RM (1989) Applied geostatistics: Oxford University Press

Ivanov S, D'yakonov A (2019) Modern deep reinforcement learning algorithms. *arXiv preprint* 1–54. doi: https://arxiv.org/abs/1906.10025

Iyakwari S, Glass HJ, Rollinson GK, Kowalczuk PB (2016) Application of near infrared sensors to preconcentration of hydrothermally-formed copper ore. *Minerals Engineering* 85:148–167. doi: 10.1016/j.mineng.2015.10.020

Jazwinski A (1970) Stochastic process and filtering theory, Volume 141. Academic Press

Jenkins R (1999) X-ray fluorescence spectrometry. Wiley, New York

Jenkins R, Snyder RL (1996) Introduction to X-ray powder diffractometry. Wiley, New York

Jewbali A (2006) Modelling geological uncertainty for stochastic short- term production scheduling in open pit metal mines. Thesis, The University of Queensland.

Jewbali A, Dimitrakopoulos R (2011) Implementation of conditional simulation by successive residuals. *Computers and Geosciences* 37:129–142. doi: 10.1016/j.cageo.2010.04.008

Jin J, Song C, Li H, Gai K, Wang J, Zhang W (2018) Real-time bidding with multi-

agent reinforcement learning in display advertising. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management. pp 2193–2201

Johnson TB (1969) Optimum open-pit mine production. In: International Symposium on Computers and Operations Research. pp 539–562

Journel A, Huijbregts C (1978) Mining geostatistics. Academic Press, London

Journel AG (2005) Beyond covariance: The advent of multiple-point geostatistics. In: Geostatistics Banff. Quantitative Geology and Geostatistics. Springer, pp 225–233

Journel AG (1974) Geostatistics for conditional simulation of ore bodies. *Economic Geology* 69:673–687

Journel AG (1994) Modeling uncertainty: Some conceptual thoughts. In: Geostatistics for the Next Century. Quantitative Geology and Geostatistics. Springer, pp 30–43

Journel AG (2007) Roadblocks to the evaluation of ore reserves—the simulation overpass and putting more geology into numerical models of deposits. In: Orebody modeling and strategic mine planning. AusIIMM, Melbourn, pp 29–32

Journel AG, Alabert FG (1990) New method for reservoir mapping. *Journal of Petroleum Technology* 42:212–218. doi: 10.2118/18324-pa

Kalman RE (1960) A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82:35–45

Kalman RE, Bucy RS (1961) New results in linear filtering and prediction theory. *Journal of Basic Engineering* 83:95. doi: 10.1115/1.3658902

Kargupta H, Srakar K, Gilligan M (2010) MineFleet ®: An overview of a widely adopted distributed vehicle performance data mining system. In: Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining. pp 37–46

Karimi M, Dehghani A, Nezamalhosseini A, Talebi S (2010) Prediction of hydrocyclone performance using artificial neural networks. *Journal of the Southern African Institute of Mining and Metallurgy* 110:207–212

Kern M, Tusa L, Leißner T, van den Boogaart KG, Gutzmer J (2019) Optimal sensor

selection for sensor-based sorting based on automated mineralogy data. *Journal of Cleaner Production* 234:1144–1152. doi: 10.1016/j.jclepro.2019.06.259

Khan A, Asad MWA (2020) A mathematical programming model for optimal cut-off grade policy in open pit mining operations with multiple processing streams. *International Journal of Mining, Reclamation and Environment* 34:149–158. doi: 10.1080/17480930.2018.1532865

Khandelwal M, Monjezi M (2013a) Prediction of flyrock in open pit blasting operation using machine learning method. *International Journal of Mining Science and Technology* 23:313–316. doi: 10.1016/j.ijmst.2013.05.005

Khandelwal M, Monjezi M (2013b) Prediction of backbreak in open-pit blasting operations using the machine learning method. *Rock Mechanics and Rock Engineering* 46:389–396. doi: 10.1007/s00603-012-0269-3

Kingma DP, Ba JL (2015) Adam: A method for stochastic optimization. *arXiv preprint* 1–15. doi: https://arxiv.org/abs/1412.6980

Kober J, Bagnell JA, Peters J (2013) Reinforcement learning in robotics : A survey. *International Journal of Robotics Research* 32:1238–1274

Koellner WG, Brown GM, Rodríguez J, Pontt J, Cortés P, Miranda H (2004) Recent advances in mining haul trucks. *IEEE Transactions on Industrial Electronics* 51:321–329. doi: 10.1109/TIE.2004.825263

Koryagin M, Voronov A (2017) Improving the organization of the shovel-truck systems in open-pit coal mines. *Transport Problems* 12:113–122. doi: 10.20858/tp.2017.12.2.11

Kozan E, Liu SQ (2016) A new open-pit multi-stage mine production timetabling model for drilling, blasting and excavating operations. *Mining Technology* 125:47–53. doi: 10.1179/1743286315Y.0000000031

Kumar D, Srinivasan S (2019) Ensemble-based assimilation of nonlinearly related dynamic data in reservoir models exhibiting non-Gaussian characteristics. *Mathematical Geosciences* 51:75–107. doi: 10.1007/s11004-018-9762-x

Kumral M (2011) Incorporating geo-metallurgical information into mine production scheduling. *Journal of the Operational Research Society* 62:60–68. doi: 10.1057/jors.2009.174

Kumral M, Dowd P (2002) Short-term mine production scheduling for industrial

minerals using multi-objective simulated annealing. In: Application of Computers and Operations Research in the Mineral Industry. pp 731–741

L'Heureux G, Gamache M, Soumis F (2013) Mixed integer programming model for short term planning in open-pit mines. *Mining Technology* 122:101–109. doi: 10.1179/1743286313Y.0000000037

Lamghari A (2017) Mine planning and oil field development: A survey and research potentials. *Mathematical Geosciences* 49:395–437. doi: 10.1007/s11004-017-9676-z

Lamghari A, Dimitrakopoulos R (2012) A diversified Tabu search approach for the open-pit mine production scheduling problem with metal uncertainty. *European Journal of Operational Research* 222:642–652. doi: 10.1016/j.ejor.2012.05.029

Lamghari A, Dimitrakopoulos R (2016a) Progressive hedging applied as a metaheuristic to schedule production in open-pit mines accounting for reserve uncertainty. *European Journal of Operational Research* 253:843–855. doi: 10.1016/j.ejor.2016.03.007

Lamghari A, Dimitrakopoulos R (2016b) Network-flow based algorithms for scheduling production in multi-processor open-pit mines accounting for metal uncertainty. *European Journal of Operational Research* 250:273–290. doi: 10.1016/j.ejor.2015.08.051

Lamghari A, Dimitrakopoulos R (2018) Hyper-heuristic approaches for strategic mine planning under uncertainty. *Computers and Operations Research* 1–18. doi: 10.1016/j.cor.2018.11.010

Lamghari A, Dimitrakopoulos R, Ferland JA (2015) A hybrid method based on linear programming and variable neighborhood descent for scheduling production in open-pit mines. *Journal of Global Optimization* 63:555–582. doi: 10.1007/s10898-014-0185-z

Lamghari A, Dimitrakopoulos R, Ferland JA (2014) A variable neighbourhood descent algorithm for the open-pit mine production scheduling problem with metal uncertainty. *Journal of the Operational Research Society* 65:1305–1314. doi: 10.1057/jors.2013.81

Lane KF (1988) The economic definition of ore: Cut-off grades in theory and practice. Mining Journal Books Limited, London

Lane KF (1984) Cutoff grades for two minerals. In: Proceedings of the 18th international symposium on application of computers and operations research in mineral the industries. pp 485–492

Larkin P (2011) Infrared and Raman spectroscopy: Principles and spectral interpretation. Elsevier, 9780123869845

Leal LAE, Westerlund M, Chapman A (2019) Autonomous industrial management via reinforcement learning: Self-learning agents for decision-making - A review. *arXiv preprint* 1–15. doi: http://arxiv.org/abs/1910.08942

Leite A, Dimitrakopoulos R (2007) Stochastic optimisation model for open pit mine planning: application and risk analysis at copper deposit. *Mining Technology* 116:109–118. doi: 10.1179/174328607X228848

Lerchs H, Grossman F (1965) Optimum design of open-pit mines. In: CIM Transactions. pp 17–24

Lessard J, De Bakker J, McHugh L (2014) Development of ore sorting and its impact on mineral processing economics. *Minerals Engineering* 65:88–97. doi: 10.1016/j.mineng.2014.05.019

Levine S, Finn C, Darrell T, Abbeel P (2016) End-to-end training of deep visuomotor policies. *Journal of Machine Learning Research* 17:1–40

Levinson Z, Dimitrakopoulos R (2019) Simultaneous stochastic optimisation of an open-pit gold mining complex with waste management. *International Journal of Mining, Reclamation and Environment* 1–15. doi: 10.1080/17480930.2019.1621441

Li C, Qiu M, Li C (2019) Deep reinforcement learning for cybersecurity. *arXiv preprint* 1–16. doi: https://arxiv.org/abs/1906.05799

Li S, Sari YA, Kumral M (2020) Optimization of mining–mineral processing integration using unsupervised machine learning algorithms. *Natural Resources Research* 1–12. doi: 10.1007/s11053-020-09628-0

Li X, Huang T, Lu DT, Niu C (2014) Accelerating experimental high-order spatial statistics calculations using GPUs. *Computers and Geosciences* 70:128–137. doi: 10.1016/j.cageo.2014.05.012

Li Z (1990) A methodology for the optimum control of shovel and truck operations in open-pit mining. *Mining Science and Technology* 10:337–340. doi:

10.1016/0167-9031(90)90543-2

Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, Silver D, Wierstra D (2015) Continuous control with deep reinforcement learning. *arXiv preprint* 1–14. doi: http://arxiv.org/abs/1509.02971

Lishchuk V, Lund C, Ghorbani Y (2019) Evaluation and comparison of different machine-learning methods to integrate sparse process data into a spatial model in geometallurgy. *Minerals Engineering* 134:156–165. doi: 10.1016/j.mineng.2019.01.032

Liu Y, Sun W, Durlofsky LJ (2019) A deep-learning-based geological parameterization for history matching complex models. *Mathematical Geosciences* 51:725–766. doi: 10.1007/s11004-019-09794-9

Luong NC, Hoang DT, Gong S, Niyato D, Wang P, Liang Y-C, Kim DI (2019) Applications of deep reinforcement learning in communications and networking: A survey. *IEEE Communications Surveys & Tutorials* 1–43. doi: 10.1109/comst.2019.2916583

Macfarlane AS, Williams TP (2014) Optimizing value on a copper mine by adopting a geometallurgical solution. *Journal of the Southern African Institute of Mining and Metallurgy* 114:929–936

Madsen H, Cañizares R (1999) Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling. *International Journal for Numerical Methods in Fluids* 31:961–981. doi: 10.1002/(SICI)1097-0363(19991130)31:6<961::AID-FLD907>3.0.CO;2-0

Mai NL, Topal E, Erten O, Sommerville B (2019) A new risk-based optimisation method for the iron ore production scheduling using stochastic integer programming. *Resources Policy* 62:571–579. doi: 10.1016/j.resourpol.2018.11.004

Mann C, Wilke FL (1992) Open pit short-term mine planning for grade control - A combination of CAD techniques and linear programming. In: Application of Computers and Operations Research in the Mineral Industry. pp 487–497

Mao H, Alizadeh M, Menache I, Kandula S (2016) Resource management with deep reinforcement learning. In: ACM Workshop on Hot Topics in Networks. pp 50–56

Mao S, Journel A (1999a) Generation of a reference petrophysical/seismic data set: the Stanford V reservoir. Stanford Cent. Reserv. Forecast. Report, Stanford, CA

Mao S, Journel AG (1999b) Conditional 3D simulation of lithofacies with 2D seismic data. *Computers and Geosciences* 25:845–862. doi: 10.1016/S0098-3004(99)00006-0

Mariethoz G, Renard P (2010) Reconstruction of incomplete data sets or images using direct sampling. *Mathematical Geosciences* 42:245–268. doi: 10.1007/s11004-010-9270-0

Matamoros MEV, Dimitrakopoulos R (2016) Stochastic short-term mine production schedule accounting for fleet allocation, operational considerations and blending restrictions. *European Journal of Operational Research* 255:911–921. doi: 10.1016/j.ejor.2016.05.050

Menabde M, Froyland G, Stone P, Yeates GA (2007) Mining schedule optimisation for conditionally simulated orebodies. In: Orebody Modelling and Strategic Mine Planning. pp 379–384

Miller RN, Carter EF, Blue ST (1999) Data assimilation into nonlinear stochastic models. *Tellus, Series A: Dynamic Meteorology and Oceanography* 51:167–194. doi: 10.3402/tellusa.v51i2.12315

Minniakhmetov I, Dimitrakopoulos R (2017a) Joint high-order simulation of spatially correlated variables using high-order spatial statistics. *Mathematical Geosciences* 49:39–66. doi: 10.1007/s11004-016-9662-x

Minniakhmetov I, Dimitrakopoulos R (2017b) A high-order, data-driven framework for joint simulation of categorical variables. In: Geostatistics Valencia. pp 287–301

Minniakhmetov I, Dimitrakopoulos R, Godoy M (2018) High-order spatial simulation using Legendre-like orthogonal splines. *Mathematical Geosciences* 50:753–780. doi: 10.1007/s11004-018-9741-2

Mitchell HL, Houtekamer PL, Pellerin G (2002) Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Monthly Weather Review* 130:2791–2808. doi: 10.1175/1520-0493(2002)130<2791:ESBAME>2.0.CO;2

Mnih V, Badia AP, Mirza L, Graves A, Harley T, Lillicrap TP, Silver D, Kavukcuoglu K (2016) Asynchronous methods for deep reinforcement learning. In:

International Conference on Machine Learning. pp 1928–1937

Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M (2013) Playing Atari with deep reinforcement learning. *arXiv preprint* 1–9. doi: http://arxiv.org/abs/1312.5602

Monjezi M, Hasanipanah M, Khandelwal M (2013) Evaluation and prediction of blast-induced ground vibration at Shur River Dam , Iran , by artificial neural network. *Neural Computing and Applications* 22:1637–1643. doi: 10.1007/s00521-012-0856-y

Montiel L, Dimitrakopoulos R (2015) Optimizing mining complexes with multiple processing and transportation alternatives: An uncertainty-based approach. *European Journal of Operational Research* 247:166–178. doi: 10.1016/j.ejor.2015.05.002

Montiel L, Dimitrakopoulos R (2017) A heuristic approach for the stochastic optimization of mine production schedules. *Journal of Heuristics* 23:397–415. doi: 10.1007/s10732-017-9349-6

Montiel L, Dimitrakopoulos R (2018) Simultaneous stochastic: optimization of production scheduling at Twin Creeks mining complex, Nevada. *Mining Enginnering* 70:48–56

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

Moradi Afrapoli A, Askari-Nasab H (2017) Mining fleet management systems: a review of models and algorithms. *International Journal of Mining, Reclamation and Environment* 33:42–60. doi: 10.1080/17480930.2017.1336607

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* 9:108. doi: 10.3390/min9020108

Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66:247–259. doi: 10.1016/j.isprsjprs.2010.11.001

Mousavi A, Kozan E, Liu SQ (2016a) Open-pit block sequencing optimization: A mathematical model and solution technique. *Engineering Optimization* 48:1932–

1950. doi: 10.1080/0305215X.2016.1142080

Mousavi A, Kozan E, Liu SQ (2016b) Comparative analysis of three metaheuristics for short-term open pit block sequencing. *Journal of Heuristics* 22:301–329. doi: 10.1007/s10732-016-9311-z

Mustapha H, Dimitrakopoulos R (2010a) A new approach for geological pattern recognition using high-order spatial cumulants. *Computers and Geosciences* 36:313–334. doi: 10.1016/j.cageo.2009.04.015

Mustapha H, Dimitrakopoulos R (2010b) High-order stochastic simulation of complex Spatially distributed natural phenomena. *Mathematical Geosciences* 42:457–485. doi: 10.1007/s11004-010-9291-8

Mustapha H, Dimitrakopoulos R (2011) HOSIM: A high-order stochastic simulation algorithm for generating three-dimensional complex geological patterns. *Computers and Geosciences* 37:1242–1253. doi: 10.1016/j.cageo.2010.09.007

Mustapha H, Dimitrakopoulos R, Chatterjee S (2011) Geologic heterogeneity representation using high-order spatial cumulants for subsurface flow and transport simulations. *Water Resources Research* 47:1–16. doi: 10.1029/2010WR009515

Myrzabekova D, Dudkin M, Młyńczak M, Muzdybayeva A, Muzdybayev M (2020) Concept of preventive maintenance in the operation of mining transportation machines. In: Advances in Intelligent Systems and Computing. Springer Verlag, pp 349–357

Nair V, Hinton GE (2010) Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th international conference on machine learning. pp 807–814

Naraghi ME, Srinivasan S (2015) Integration of seismic and well data to characterize facies variation in a carbonate reservoir-the tau model revisited. In: Petroleum Geostatistics. European Association of Geoscientists and Engineers, EAGE, pp 243–247

Navarra A, Grammatikopoulos T, Waters K (2018) Incorporation of geometallurgical modelling into long-term production planning. *Minerals Engineering* 120:118–126. doi: 10.1016/j.mineng.2018.02.010

Neves J, Pereira MJ, Pacheco N, Soares A (2018) Updating mining resources with

uncertain data. *Mathematical Geosciences* 51:905–924. doi: 10.1007/s11004-018-9759-5

Nguyen D, Bui X (2015) A real-time regulation model in multi-agent decision support system for open pit mining. In: International Symposium Continuous Surface Mining-Aachen. Springer, Cham, pp 255–262

Oliver D, Reynolds A, Liu N (2008) Inverse theory for petroleum reservoir characterization and history matching. Cambridge University Press, Cambridge

Oliver DS (1996) Multiple realizations of the permeability field from well test data. *SPE Journal* 1:145–154

Oliver DS, Chen Y (2011) Recent progress on reservoir history matching: A review. *Computational Geosciences* 15:185–221. doi: 10.1007/s10596-010-9194-2

Oliver DS, Cunha LB, Reynolds AC (1997) Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Mathematical Geology* 29:61–91. doi: 10.1007/BF02769620

Ortiz JM, Kracht W, Pamparana G, Haas J (2019) Optimization of a SAG mill energy system: Integrating rock hardness, solar irradiation, climate change, and demand-side management. *Mathematical Geosciences* 1–25. doi: 10.1007/s11004-019-09816-6

Ozdemir B, Kumral M (2019) Simulation-based optimization of truck-shovel material handling systems in multi-pit surface mines. *Simulation Modelling Practice and Theory* 95:36–48. doi: 10.1016/j.simpat.2019.04.006

Paduraru C, Dimitrakopoulos R (2019) Responding to new information in a mining complex: Fast mechanisms using machine learning. *Mining Technology* 128:129–142. doi: 10.1080/25726668.2019.1577596

Paduraru C, Dimitrakopoulos R (2018) Adaptive policies for short-term material flow optimization in a mining complex. *Mining Technology* 127:56–63. doi: 10.1080/14749009.2017.1341142

Paithankar A, Chatterjee S (2019) Open pit mine production schedule optimization using a hybrid of maximum-flow and genetic algorithms. *Applied Soft Computing Journal* 81:105507. doi: 10.1016/j.asoc.2019.105507

Paithankar A, Chatterjee S, Goodfellow R, Asad MWA (2020) Simultaneous stochastic optimization of production sequence and dynamic cut-off grades in an

open pit mining operation. *Resources Policy* 66:101634. doi: 10.1016/j.resourpol.2020.101634

Panda L, Sahoo AK, Tripathy A, Biswal SK, Sahu AK (2012) Application of artificial neural network to study the performance of jig for beneficiation of non-coking coal. *Fuel* 97:151–156. doi: 10.1016/j.fuel.2012.02.018

Panda L, Tripathy SK (2014) Performance prediction of gravity concentrator by using artificial neural network-a case study. *International Journal of Mining Science and Technology* 24:461–465. doi: 10.1016/j.ijmst.2014.05.007

Panzeri M, Della Rossa EL, Dovera L, Riva M, Guadagnini A (2016) Integration of Markov mesh models and data assimilation techniques in complex reservoirs. *Computational Geosciences* 20:637–653. doi: 10.1007/s10596-015-9540-5

Patel AK, Chatterjee S (2016) Computer vision-based limestone rock-type classification using probabilistic neural network. *Geoscience Frontiers* 7:53–60. doi: 10.1016/j.gsf.2014.10.005

Petropoulos GP, Partsinevelos P, Mitraka Z (2013) Change detection of surface mining activity and reclamation based on a machine learning approach of multi-temporal Landsat TM imagery. *Geocarto International* 28:323–342. doi: 10.1080/10106049.2012.706648

Pimentel BS, Mateus GR, Almeida FA (2010) Mathematical models for optimizing the global mining supply chain. *Intelligent Systems in Operations* 133–163. doi: 10.4018/978-1-61520-605-6.ch008

Place D, Whittle G, Baxter N (2018) GEOVIA Whittle simultaneous optimization

Quigley M, Dimitrakopoulos R (2019) Incorporating geological and equipment performance uncertainty while optimizing short-term mine production schedules. *International Journal of Mining Reclamation and Environment* 1–22. doi: 10.1080/17480930.2019.1658923

Rajabinasab B, Asghari O (2019) Geometallurgical domaining by cluster analysis: Iron ore deposit case study. *Natural Resources Research* 28:665–684. doi: 10.1007/s11053-018-9411-6

Ramazan S, Dimitrakopoulos R (2013) Production scheduling with uncertain supply: A new solution to the open pit mining problem. *Optimization and Engineering* 14:361–380. doi: 10.1007/s11081-012-9186-2

Ramazan S, Dimitrakopoulos R (2007) Stochastic optimisation of long-term production scheduling for open pit mines with a new integer programming formulation. In: Orebody Modelling and Strategic Mine Planning. pp 385–391

Ramírez-Gallego S, Fernández A, García S, Chen M, Herrera F (2018) Big Data: Tutorial and guidelines on information and process fusion for analytics algorithms with MapReduce. *Information Fusion* 42:51–61. doi: 10.1016/j.inffus.2017.10.001

Ravenscroft PJ (1992) Risk analysis for mine scheduling by conditional simulation. *Transactions of the Institution of Mining and Metallurgy Section A Mining Industry* 101:104–108. doi: 10.1016/0148-9062(93)90969-K

Remy N (2005) The Stanford geostatistical modeling software: A tool for new algorithms development. In: Geostatistics Banff. Springer, Dordrecht, pp 865–871

Remy N, Boucher A, Wu J (2009) Applied geostatistics with SGeMS: A user's guide. Cambridge University Press

Rendu J-M (2014) An introduction to cut-off grade estimation. Society for Mining, Metallurgy & Exploration, Englewood, Colarado

Rosa L, David, Valery W, Wortley M, Ozkocak T, Pike M (2007) The use of radio frequency ID tags to track ore in mining operations. In: Proceedings of the 33rd application of computers and operations research in the mineral Industries. pp 601–606

Rossi ME, Deutsch C V. (2013) Mineral resource estimation. Springer Science & Business Media, New York

Ruder S (2016) An overview of gradient descent optimization algorithms. *arXiv preprint* 1–14. doi: https://arxiv.org/abs/1609.04747

Saliba Z, Dimitrakopoulos R (2019a) Simultaneous stochastic optimization of an open pit gold mining complex with supply and market uncertainty. *Mining Technology* 128:216–229. doi: 10.1080/25726668.2019.1626169

Saliba Z, Dimitrakopoulos R (2019b) An application of simultaneous stochastic optimisation of an open-pit mining complex with tailings management. *International Journal of Mining, Reclamation and Environment* 1–16. doi: 10.1080/17480930.2019.1688954

Sambridge M (1999) Geophysical inversion with a neighbourhood algorithm--II. Appraising the ensemble. *Geophysical Journal International* 138:727–746. doi: 10.1046/j.1365-246X.1999.00900.x

Sarma P, Durlofsky LJ, Aziz K, Chen WH (2006) Efficient real-time reservoir management using adjoint-based optimal control and model updating. *Computational Geosciences* 10:3–36. doi: 10.1007/s10596-005-9009-z

Sato Y (2019) Model-free reinforcement learning for financial portfolios: A brief survey. *arXiv preprint* 1–20. doi: http://arxiv.org/abs/1904.04973

Schleifer J (1996) Short-range planning - from scheduling to sequencing. In: Application of Computers and Operations Research in the Mineral Industry. pp 183–188

Schrittwieser J, Antonoglou I, Hubert T, Simonyan K, Sifre L, Guez A, Lockhart E, Hassabis D, Graepel T, Lillicrap T, Silver D (2019) Mastering Atari, Go, Chess and Shogi by planning with a learned model. *arXiv preprint* 1–21. doi: https://arxiv.org/abs/1911.08265

Schulze-Riegert R, Ghedan S (2007) Modern techniques for history matching. In: International Forum on Reservoir Simulation. pp 9–13

Sepúlveda E, Dowd PA, Xu C (2018) Fuzzy clustering with spatial correction and its application to geometallurgical domaining. *Mathematical Geosciences* 50:895–928. doi: 10.1007/s11004-018-9751-0

Sepulveda E, Dowd PA, Xu C, Addo E (2017) Multivariate modelling of geometallurgical variables by projection pursuit. *Mathematical Geosciences* 49:121–143. doi: 10.1007/s11004-016-9660-z

Shariati H, Yeraliyev A, Terai B, Tafazoli S, Ramezani M (2019) Towards autonomous mining via intelligent excavators. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp 26–32

Shirangi MG (2017) Advanced techniques for closed-loop reservoir optimization under uncertainty (Doctoral dissertation). Thesis, Stanford University, Stanford, California.

Shuai Y, White C, Sun T, Feng Y (2016) A gathered EnKF for continuous reservoir model updating. *Journal of Petroleum Science and Engineering* 139:205–218. doi: 10.1016/j.petrol.2016.01.005

Siami-Irdemoosa E, Dindarloo SR (2015) Prediction of fuel consumption of mining dump trucks: A neural networks approach. *Applied Energy* 151:77–84. doi: 10.1016/j.apenergy.2015.04.064

Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 529:484–489. doi: 10.1038/nature16961

Silver D, Hubert T, Schrittwieser J, Antonoglou I, Graepel T, Lillicrap T, Simonyan K, Hassabis D, Turing A, Shannon C (2018) A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362:1140–1144

Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y, Lillicrap T, Hui F, Sifre L, Van Den Driessche G, Graepel T, Hassabis D (2017) Mastering the game of Go without human knowledge. *Nature* 550:354–359. doi: 10.1038/nature24270

Singh JP, Thakur SN (2007) Laser-induced breakdown spectroscopy. Elsevier, 9780444517340

Skjervheim JA, Evensen G, Aanonsen SI, Ruud BO, Johansen TA (2007) Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter. *SPE Journal* 12:282–292. doi: 10.2118/95789-PA

Sládková D, Kapica R, Vrubel M (2011) Global navigation satellite system (GNSS) technology for automation of surface mining. *International Journal of Mining, Reclamation and Environment* 25:284–294. doi: 10.1080/17480930.2011.608879

Smith ML (1998) Optimizing short-term production schedules in surface mining: Integrating mine modeling software with AMPL/CPLEX. *International Journal of Surface Mining, Reclamation and Environment* 12:149–155. doi: 10.1080/09208118908944038

Soares A (2001) Direct sequential simulation and cosimulation. *Mathematical Geology* 33:911–925. doi: 10.1023/A:1012246006212

Soares A, Nunes R, Azevedo L (2017) Integration of uncertain data in geostatistical

modelling. *Mathematical Geosciences* 49:253–273. doi: 10.1007/s11004-016-9667-5

Spleit M, Dimitrakopoulos R (2017) Risk management and long-term production schedule optimization at the LabMag iron ore deposit in Labrador, Canada. *Mining Enginnering* 69:47–53

Stone P, Froyland G, Menabde M, Law B, Pasyar R, Monkhouse P (2007) Blasor — Blended iron ore mine planning optimisation at Yandi. In: Orebody Modelling and Strategic Mine Planning. pp 1–7

Straubhaar J, Renard P, Mariethoz G, Froidevaux R, Besson O (2011) An improved parallel multiple-point algorithm using a list approach. *Mathematical Geosciences* 43:305–328. doi: 10.1007/s11004-011-9328-7

Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology* 34:1–21. doi: 10.1023/A:1014009426274

Strebelle S, Cavelius C (2014) Solving speed and memory issues in multiple-point statistics simulation program SNESIM. *Mathematical Geosciences* 46:171–186. doi: 10.1007/s11004-013-9489-7

Sutton R, Barto A (2017) Reinforcement learning: An introduction, Second edi. MIT press

Sutton RS, McAllester D, Singh S, Mansour Y (2000) Policy gradient methods for reinforcement learning with function approximation. In: Proceedings of advances in neural information processing systems. pp 1057–1063

Switzer P, Green AA (1984) Min/Max autocorrelation factors for multivariate spatial imagery. Tech. Rep. No. 6, Dep. Stat. Stanford Univ. 14

Ta CH, Ingolfsson A, Doucette J (2013) A linear model for surface mining haul truck allocation incorporating shovel idle probabilities. *European Journal of Operational Research* 231:770–778. doi: 10.1016/j.ejor.2013.06.016

Ta CH, Kresta J V., Forbes JF, Marquez HJ (2005) A stochastic optimization approach to mine truck allocation. *International Journal of Surface Mining, Reclamation and Environment* 19:162–175. doi: 10.1080/13895260500128914

Tahmasebi P, Sahimi M (2015) Geostatistical simulation and reconstruction of porous media by a cross-correlation function and integration of hard and soft data.

Transport in Porous Media 107:871–905. doi: 10.1007/s11242-015-0471-3

Tang M, Liu Y, Durlofsky LJ (2019) A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. *arXiv preprint* 1–47. doi: http://arxiv.org/abs/1908.05823

Uhlenbeck GE, Ornstein LS (1930) On the Theory of the Brownian Motion. *Physical Review* 36:823. doi: 10.1143/JPSJ.13.1266

Underwood R, Tolwinski B (1998) A mathematical programming viewpoint for solving the ultimate pit problem. *European Journal of Operational Research* 107:96–107. doi: 10.1016/S0377-2217(97)00141-0

Upadhyay SP, Askari-Nasab H (2016) Truck-shovel allocation optimisation: A goal programming approach. *Mining Technology* 125:82–92. doi: 10.1179/1743286315Y.0000000024

Upadhyay SP, Askari-Nasab H (2018) Simulation and optimization approach for uncertainty-based short-term planning in open pit mines. *International Journal of Mining Science and Technology* 28:153–166. doi: 10.1016/j.ijmst.2017.12.003

Van Loon M, Builtjes PJH, Segers AJ (2000) Data assimilation of ozone in the atmospheric transport chemistry model LOTOS. In: Environmental Modelling and Software. Elsevier Science Ltd, pp 603–609

Vargas-Guzmán JA, Dimitrakopoulos R (2003) Successive nonparametric estimation of conditional distributions. *Mathematical Geology* 35:39–52. doi: 10.1023/A:1022361028297

Vargas-Guzmán JA, Dimitrakopoulos R (2002) Conditional simulation of random fields by successive residuals. *Mathematical Geology* 34:597–611. doi: 10.1023/A:1016099029432

Vasylchuk Y V., Deutsch C V. (2018) Improved grade control in open pit mines. *Mining Technology* 127:84–91. doi: 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* 37:451–475. doi: 10.1007/s11004-005-6660-9

Verly GW (1993) Sequential Gaussian cosimulation: A simulation method integrating several types of information. In: Geostatistics Tróia. pp 543–554

Verret F-O, Chiasson G, McKen A (2011) Sag mill testing - An overview of the test procedures available To characterize ore grindability. In: Proceedings of International Autogenous and Semiautogenous Grinding and High Pressure Grinding Roll Technology. pp 1–14

Vo HX, Durlofsky LJ (2014) A new differentiable parameterization based on principal component analysis for the low-dimensional representation of complex geological models. *Mathematical Geosciences* 46:775–813. doi: 10.1007/s11004-014-9541-2

Wambeke T, Benndorf J (2017) A Simulation-based geostatistical approach to real-time reconciliation of the grade control model. *Mathematical Geosciences* 49:1–37. doi: 10.1007/s11004-016-9658-6

Wambeke T, Benndorf J (2018) A study of the influence of measurement volume ,blending ratios and sensor precision on real-time reconciliation of grade control models. *Mathematical Geosciences* 50:801–826. doi: 10.1007/s11004-018-9740-3

Wambeke T, Elder D, Miller A, Benndorf J, Peattie R (2018) Real-time reconciliation of a geometallurgical model based on ball mill performance measurements – A pilot study at the Tropicana gold mine. *Mining Technology* 127:115–130. doi: 10.1080/25726668.2018.1436957

Wei Q, Zhu S, Du C (2011) Study on key technologies of Internet of Things perceiving mine. *Procedia Engineering* 26:2326–2333. doi: 10.1016/j.proeng.2011.11.2442

Wen XH, Chen WH (2005) Real-time reservoir model updating using ensemble Kalman filter. In: SPE Reservoir Simulation Symposium. pp 167–180

Wen XH, Chen WH (2006) Real-time reservoir model updating using ensemble Kalman filter with confirming option. *SPE Journal* 11:431–442. doi: 10.2118/92991-PA

Whittle G (2007) Global asset optimisation. In: Orebody Modelling and Strategic Mine Planning. pp 331–336

Whittle J (2010) The global Optimiser works – What next ? In: Orebody Modelling and Strategic Mine Planning. pp 3–5

Whittle J (2014) Not for the faint-hearted. In: Orebody Modelling and Strategic Mine

Planning. pp 3–5

Whittle J, Whittle G (2007) Global long-term optimisation of very large mining complexes. In: Application of Computers and Operations Research in the Mineral Industry. pp 253–260

Wilke FL, Reimer TH (1977) Optimizing the short term production schedule for an open pit iron ore mining operation. In: Application of Computers and Operations Research in Mineral Industry. pp 425–433

Williams SR, Richardson JM (2004) Geometallurgical mapping: A new approach that reduces technical risk. In: Annual Meeting of the Canadian Mineral Processors. pp 241–268

Xu T, Gómez-Hernández JJ (2015) Inverse sequential simulation: Performance and implementation details. *Advances in Water Resources* 86:311–326. doi: 10.1016/j.advwatres.2015.04.015

Xu T, Hernández JG (2019) Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble Kalman filter. *Advances in Water Resources* 112:106–123

Xu T, Jaime Gómez-Hernández J, Li L, Zhou H (2013) Parallelized ensemble Kalman filter for hydraulic conductivity characterization. *Computers and Geosciences* 52:42–49. doi: 10.1016/j.cageo.2012.10.007

Xue L, Zhang D (2014) A multimodel data assimilation framework via the ensemble Kalman filter. *Water Resources Research* 50:4197–4219. doi: 10.1002/2013WR014525

Yao L, Dimitrakopoulos R, Gamache M (2018) A new computational model of high-order stochastic simulation based on spatial Legendre moments. *Mathematical Geosciences* 50:929–960. doi: 10.1007/s11004-018-9744-z

Yüksel C, Benndorf J, Lindig M, Lohsträter O (2017) Updating the coal quality parameters in multiple production benches based on combined material measurement: A full case study. *International Journal of Coal Science and Technology* 4:159–171. doi: 10.1007/s40789-017-0156-3

Yüksel C, Minnecker C, Shishvan MS, Benndorf J, Buxton M (2018) Value of information introduced by a resource model updating framework. *Mathematical Geosciences* 51:1–19. doi: 10.1007/s11004-018-9770-x

Yüksel C, Thielemann T, Wambeke T, Benndorf J (2016) Real-time resource model updating for improved coal quality control using online data. *International Journal of Coal Geology* 162:61–73. doi: 10.1016/j.coal.2016.05.014

Zhang J, Dimitrakopoulos R (2017) A dynamic-material-value-based decomposition method for optimizing a mineral value chain with uncertainty. *European Journal of Operational Research* 258:617–625. doi: 10.1016/j.ejor.2016.08.071

Zhang L, Xia X (2015) An integer programming approach for truck-shovel dispatching problem in open-pit mines. *Energy Procedia* 75:1779–1784. doi: 10.1016/j.egypro.2015.07.469

Zhang T, Switzer P, Journel A (2006) Filter-based classification of training image patterns for spatial simulation. *Mathematical Geology* 38:63–80. doi: 10.1007/s11004-005-9004-x

Zhao X, Xia L, Tang J, Yin D (2019) Deep reinforcement learning for search, recommendation, and online advertising: a survey. *ACM SIGWEB Newsletter* 4:1–15. doi: 10.1145/3320496.3320500

Zheng G, Zhang F, Zheng Z, Xiang Y, Yuan NJ, Xie X, Li Z (2018) DRN: A deep reinforcement learning framework for news recommendation. In: International World Wide Web Conference. pp 167–176

Zhou Z, Li X, Zare RN (2017) Optimizing chemical reactions with deep reinforcement learning. *ACS Central Science* 3:1337–1344. doi: 10.1021/acscentsci.7b00492

Zhu H (1991) Modeling mixtures of spatial distributions with integration of soft data. Thesis, Standford University.