Connecting Planning Horizons in Mining Complexes with Reinforcement Learning and Stochastic Programming: Integrating Additional Information, Preconcentration, Reclamation and Waste Management

Zachary Levinson

Department of Mining and Materials Engineering

McGill University, Montreal, Canada

August 2023

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

Abstract

A mining complex is an integrated supply chain with several components including open-pit and underground mines, crushers, preconcentration and processing facilities, transportation systems, stockpiles, waste dump facilities, and ports. Simultaneous stochastic optimization frameworks for optimizing mining complexes aim to globally optimize components together in a single mathematical programming framework that maximizes value and minimizes risk, by managing uncertainty directly in the stochastic programming formulation. Production planning and scheduling for mining complexes begins with optimizing the long-term production schedule and then progressively optimizing shorter timescales to provide the associated production and financial forecasts at each timescale. Short-term production scheduling decisions are bound by the previously determined long-term production schedule, which can compromise value and prevent major objectives from being obtained. A potential opportunity exists to jointly optimize short- and long-term production schedules, which is expected to provide significant benefits to the short- and long-term production and financial forecasts, while also improving operational compliance between timescales. This thesis develops a stochastic programming framework for jointly optimizing short- and long-term production schedules under uncertainty by connecting planning timescales with an embedded reinforcement learning agent. In addition, critical environmental management considerations are addressed to minimize the production of hazardous waste and improve reclamation practices to enhance environmental performance. Several critical components are integrated into the simultaneous stochastic optimization framework including preconcentration, waste management, and progressive reclamation. Lastly, a novel optimization framework for infill drilling is developed that links infill drilling to long-term production scheduling, as an approach to quantify the value of collecting additional information related to the material supply.

Résumé

Un complexe minier est une chaîne d'approvisionnement intégrée composée de plusieurs éléments, notamment des mines à ciel ouvert et souterraines, des concasseurs, des installations de préconcentration et de traitement, des systèmes de transport, des stocks, des décharges et des ports. Les cadres d'optimisation stochastique simultanée pour l'optimisation des complexes miniers visent à optimiser globalement tous les composants dans un cadre de programmation mathématique unique qui maximise la valeur et minimise le risque, en gérant l'incertitude directement dans la formulation de la programmation stochastique. La planification et l'ordonnancement de la production pour les complexes miniers commencent par l'optimisation du programme de production à long terme, puis par l'optimisation progressive d'échelles de temps plus courtes afin de fournir les prévisions de production et financières associées à chaque échelle de temps. Les décisions relatives à la programmation de la production à court terme sont liées à la programmation de la production à long terme déterminée précédemment, ce qui peut compromettre la valeur et empêcher d'atteindre les objectifs. Il existe une possibilité d'optimiser conjointement les programmes de production à court et à long terme, ce qui devrait apporter des avantages significatifs aux prévisions financières et de production à court et à long terme, tout en améliorant la conformité opérationnelle entre les échelles de temps. Cette thèse développe un cadre de programmation stochastique pour optimiser conjointement les calendriers de production à court et à long terme dans l'incertitude en connectant les échelles de planification avec un agent d'apprentissage par renforcement intégré. En outre, des considérations critiques de gestion environnementale sont abordées afin de minimiser la production de déchets dangereux et d'améliorer les pratiques de remise en état afin d'améliorer les performances environnementales. Plusieurs composantes essentielles sont intégrées dans le cadre d'optimisation stochastique simultanée, notamment la préconcentration, la gestion des déchets et la remise en état progressive. Enfin, un nouveau cadre d'optimisation pour le forage intercalaire est développé qui relie le forage intercalaire à la programmation de la production à long terme, en tant qu'approche pour quantifier la valeur de l'information supplémentaire et intégrer l'information supplémentaire dans les entrées des cadres d'optimisation stochastiques simultanés.

Contributions of Authors

The author of this thesis is the primary author for all manuscripts contained within. Professor Roussos Dimitrakopoulos is the supervisor of the author's Ph.D. and is included as co-author in all of the following articles. Julien Keutchayan is a postdoctoral fellow under Professor Roussos Dimitrakopoulos and is included as an author for Chapter 3.

Chapter 2 – Levinson, Z. and Dimitrakopoulos, R. (Submitted). A reinforcement learning approach for selecting infill drilling locations considering long-term production planning in mining complexes with supply uncertainty. *Mining Technology*.

Chapter 3 – Levinson, Z., Dimitrakopoulos, R., and Keutchayan, J. (2023). Simultaneous stochastic optimization of an open-pit mining complex with preconcentration using reinforcement learning. *Applied Soft Computing*, 110180. doi:10.1016/j.asoc.2023.110180

Chapter 4 – Levinson, Z. and Dimitrakopoulos, R. (Submitted). Connecting planning horizons in mining complexes with reinforcement learning and stochastic programming. *Resources Policy*.

Chapter 5 – Levinson, Z. and Dimitrakopoulos, R. (Submitted). Simultaneous stochastic optimization of mining complexes: Integrating progressive reclamation and waste management with contextual bandits. *International Journal of Mining, Reclamation and Environment*.

Acknowledgements

A special thank you is expressed to my supervisor Professor Roussos Dimitrakopoulos for his continuous support in the development of the work in this thesis. There were several times over the years where the challenges addressed seemed insurmountable. Your encouragement and continuous pressure to fulfill this degree made this possible.

I would also like to send a special thank you to Barbara for your help navigating McGill. As for the students, staff and members of the COSMO Laboratory I appreciate the opportunity we have had to learn from each other. Ashish, Christian, Cristina, Jacob, Joao, Julien, Laura, Liam, Lingqing, Will, and Yassine it has been fun. My deepest gratitude is sent to Julien for providing mentorship, inspiration and lessons that only a few would be lucky enough to receive over their lifetime. This degree would not have been nearly as interesting without your help and desire to challenge the status quo.

The COSMO Laboratory at McGill University has provided me the opportunity to travel and present at various locations around the world. The people I have met through these opportunities have shaped my research and understanding of the industry immensely. This would not be made possible without the support from our industry sponsors and the public funding agencies that understand the importance of this research.

Most importantly, words cannot express my gratitude for my incredible partner Courtney. She has been my greatest supporter and the main reason for staying grounded throughout this degree. With endless discussions regarding my work, countless edits, she truly shaped this research and helped make it presentable to all audiences. All in all, with keeping me healthy and ensuring that we still have some fun along the way. This is for you my dear.

I would also like to thank my parents, Vavo and Vovo, and my sister for allowing me to chase this opportunity and providing immense support along the way from afar. Bram, Bryce and Dan thanks for listening and dealing with my endless stories and providing a place to stay when needed. Each of you are appreciated for your part in this journey. Lastly, I thank those in industry who replied to my cold calls (Katie, Stephen and Ted) who shared with me their valuable experiences. Without people like you who take this time, the research community would not see the true impacts that they can make on industry.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Overview

A mining complex is a multifaceted supply chain that begins by extracting raw materials from one or more mineral deposits followed by the transformation of materials into marketable products through several processes. Mining complexes may include mines, preconcentration facilities, leach pads and processing plants, transportation systems, stockpiles, waste dump facilities, and ports (Pimentel et al. 2010; Montiel and Dimitrakopoulos 2015; 2017; 2018; Goodfellow and Dimitrakopoulos 2016; 2017; Dimitrakopoulos and Lamghari 2022). Understanding the mineral deposits and the key attributes that impact the production schedule and operational performance of a mining complex are the paramount drivers to its success. Uncertainty and local variability of the materials within the mineral deposits contributes major sources of risk that are shown to significantly affect the success of operations and directly impact production and financial forecasts (Ravenscroft 1992; Dowd 1994; Vallee 2000). Developing technologies that manage uncertainty and simultaneously optimize the components of a mining complex can further increase value by capitalizing on advantageous synergies and minimizing technical risk.

Long-term strategic mine planning under uncertainty aims to optimize a mining complex to maximize net present value and minimize risk. A long-term strategic mine plan is created by optimizing the production scheduling decisions within a mining complex. Advancements in modelling techniques and optimization algorithms have led to several stochastic programming approaches for simultaneously optimizing long-term production schedules in mining complexes under uncertainty (Goodfellow and Dimitrakopoulos 2016; 2017; Montiel and Dimitrakopoulos 2017; 2018; Dimitrakopoulos and Lamghari 2022). These are known as simultaneous stochastic optimization frameworks. Supply uncertainty related to the quantity and quality of the attributes of interest within the mineral deposits is incorporated with a set of equiprobable stochastic orebody simulations that are inputs to the optimization framework (Goovaerts 1997). Similarly, price uncertainty is included with stochastic price simulations (Schwartz 1997). Simultaneous stochastic optimization approaches have led to substantial improvements in long-term production schedules as showcased in several studies over the years (Kumar and Dimitrakopoulos 2019; Saliba and

Dimitrakopoulos 2019; 2020; Paithankar et al. 2020; Levinson and Dimitrakopoulos 2020b; LaRoche-Boisvert et al. 2021).

In contrast, conventional approaches for optimizing a long-term production schedule fail to consider the interactions between key components with interdependent behaviour (Dagdelen 2001; Hustrulid et al. 2013), which leads to sub-optimal production plans and associated forecasts. Early work in the area of simultaneous optimization addressed this limitation by optimizing several components together including multi-mine production schedules, stockpiling, blending, and destination policies that showed large improvements to the production schedule forecasts when compared to conventional approaches (Hoerger et al. 1999b; Stone et al. 2005; 2018; Whittle 2007; Whittle and Whittle 2007; Whittle and Burks 2010). However, these early simultaneous optimization approaches, similar to there conventional counterparts, suffer from major limitations including: (1) the material uncertainty and local variability of material is ignored which leads to substantial risk of missing production and financial forecasts and/or project failure; and (2) simplifications are taken to reduce the complexity of the optimization problem. These limitations are addressed and improved by optimizing long-term production schedules with simultaneous stochastic optimization frameworks.

Common practice for production planning begins with optimizing the yearly long-term production schedule over the life of a mining project and then progressively scheduling shorter timescales to provide operational plans (Gershon 1983). Short-term production schedules are then optimized to meet production forecasts that correspond with the yearly long-term production schedule over shorter timescales (days, weeks, or months) given the planned areas of extraction (Fytas et al. 1993; Hustrulid et al. 2013; Blom et al. 2018; 2019). A major limitation of conventional optimization and planning approaches is that they fail to consider the connected behaviour between different planning timescales. Long-term production scheduling decisions focus on maximizing net present value and ignore short-term operational impacts. Meanwhile, short-term production scheduling decision are optimized within constrained areas that were decided by the long-term production schedule, which can make it difficult to create operational short-term plans that meet long-term forecasts, potentially jeopardizing value. Similar to the idea of optimizing components together to capitalize on synergies, this thesis aims to address the limitations of optimizing the short- and long-term production schedule separately. Connecting short- and long-term production schedules is expected to increase value, mitigate risk and integrate key operational components in the production planning process for mining complexes.

Connecting production schedules with different timescales requires understanding the implications of planning the short- and long-term production schedule directly within the optimization framework. This is a significantly more challenging optimization requirement as there are far more decisions to consider. Simultaneous stochastic optimization frameworks for short-term production scheduling in mining complexes have been developed and address the previously mentioned limitations of ignoring uncertainty and independently optimizing components (Both and Dimitrakopoulos 2020; 2021; 2023a; 2023b). However, these optimization approaches still require reoptimizing the short-term production schedule each time a long-term production schedule is adjusted, which makes it challenging to develop a method for optimizing short- and long-term production schedules together. Reinforcement learning frameworks can be used to directly compute policies by learning associations between situations and actions given the rewards or punishments that are received (Sutton and Barto 2018). For optimizing production scheduling decisions, reinforcement learning has been applied to compute policies for short-term production schedules that considers updated production data (Paduraru and Dimitrakopoulos 2018; 2019; Kumar et al. 2020; Kumar and Dimitrakopoulos 2021). Different from the simultaneous stochastic optimization approaches, learning based methods are used to find policies for adapting short-term production scheduling decisions that can be reused for re-optimization of the entire mining complex by mapping unseen situations to actions. Previously learned policies can then be applied to optimize the short-term production schedule and evaluate the impact on the long-term production schedule. This has not been explored previously and is investigated in this thesis.

Furthermore, global mining companies continually aim to advance technologies that improve operational and environmental performance. Trends across different commodities show that the quality of material that is available for extraction is generally decreasing, which can cause critical issues related to energy and water consumption requirements in the mining industry (Mudd 2007; 2012; Northey et al. 2014). As a result of lower grade material entering the processing facilities, there is higher energy and water usage per unit of product produced which leads to higher production costs. Preconcentration technologies separate valuable material from waste by leveraging unique material characteristics. Examples of these technologies include size, density, and sensor based sorting of ore and waste materials that can be placed in different parts of the mining complex (Adair et al. 2020). These technologies provide opportunities to reduce energy and water requirements, thereby decreasing costs and improving feed grades at the processing facilities. Several examples show the benefits of separating waste material from the processing feed prior to energy intensive crushing and grinding (Burns and Grimes 1986; Bowman and Bearman 2014). For the complete value of these technologies to be realized, the benefits must be quantified when optimizing the short- and long-term production schedules. Materials with beneficial preconcentration qualities can provide improvements to production plans by optimizing components together. Studies have been completed using conventional optimization approaches to assess the value of preconcentration (Scott 2013; Espejel et al. 2017). However, these changes are local and fail to consider the connections between the extraction sequence, destination policies and preconcentration decisions. Fathollahzadeh et al. (2021) demonstrate the benefits of jointly optimizing preconcentration decisions with the long-term production schedule but ignore supply uncertainty. This thesis aims to develop a framework for jointly optimizing short-term production schedules to assess the impact of preconcentration at shorter timescales, where it is expected to be more beneficial due to the enhanced operational flexibility. Additionally, uncertainty related to the material supply and preconcentration decisions are incorporated into the simultaneous stochastic optimization framework for mining complexes.

Failing to consider waste management and reclamation, two crucial components related to production planning and optimization, can adversely affect a mining complex (McHaina 2001). These considerations are often considered after the optimized long-term production schedule is completed, which eliminates flexibility in the planning decisions to assist in the environmental management of waste materials and subsequent reclamation of waste dump facilities. Waste management is critical for ensuring the safe containment of waste materials within the waste dump facilities (Hustrulid et al. 2013). In certain mineral deposits, the waste rock characteristics can lead to acid rock drainage and associated metal leaching, a large environmental concern (Price 2003; Nordstrom 2011). Acid rock drainage can be tackled by blending waste materials with different geochemical properties to create a non-acid generating waste rock product for safe placement in waste dump facilities (Mehling et al. 1997; Day 2022). Furthermore, reclamation and closure requirements are also critical and require a strategy to improve environmental management and reduce environmental risks (Straker et al. 2020). Progressive reclamation is one option that is carried out in parallel to production and provides advantages as information and data can be collected to improve reclamation practices. In addition, there are major benefits to directly placing material for reclamation activities including reduced rehandle and improved soil development (Macdonald et al. 2015). In this thesis, these critical components are integrated into the simultaneous stochastic optimization framework for mining complexes and provide potential to improve environmental practices that alleviate long-term risk and increase the confidence of local governments and stakeholders in mining projects.

Lastly, simultaneous stochastic optimization frameworks are driven by their inputs, a set of stochastic orebody simulations that describe the uncertainty and local variability of the attributes within the mineral deposits. The extraction, transportation and transformation of raw materials into valuable products is based on the understanding of the materials that are available to be extracted. A critical management decision in a mining complex is determining whether to collect additional information to update the stochastic orebody simulations. Infill drilling provides information that can be used to update the stochastic orebody simulations and the associated optimized production scheduling decisions. Previous frameworks utilize kriging variance to select infill drilling locations, which is influenced by the spacing between drillholes and ignores the influence of uncertainty and local variability (Diehl and David 1982; Gershon et al. 1988; Soltani and Safa 2015). Stochastic simulations of infill drilling data have also been used to assess the impact of different drillhole configurations by comparing the classification of materials and economic indicators (Boucher et al. 2005; Dirkx and Dimitrakopoulos 2018). The limitation of these approaches is that they fail to relate the impact of drilling to production scheduling decisions and ignore the uncertainty of the materials in the ground. The information gathered through infill drilling may alter the optimized long-term production schedule, which can provide significant value and improve production guidance. An infill drilling optimization framework is developed in this thesis to link infill drilling to long-term production using a newly developed criterion.

1.2 Literature Review

The following sections provide a critical review of the research topics discussed in this thesis. Section 1.2.1 provides an overview of the main developments that consider simultaneous stochastic optimization for long-term production scheduling. Section 1.2.2 reviews advancement in preconcentration technologies and integration into long-term production scheduling. Section 1.2.3 reviews the main developments in short-term production scheduling and includes the advancement that led to the simultaneous stochastic optimization of short-term production schedules. Section 1.2.4 then discusses advancements for short-term production scheduling using adaptive policies. Finally, Section 1.2.5 discusses methods for modelling supply uncertainty and updating stochastic orebody simulations in mining complexes.

1.2.1 Simultaneous stochastic optimization for long-term production scheduling

Strategic or long-term mine planning is completed to optimize the yearly production schedule of a mining complex over its lifetime. The ultimate goal is to determine a production schedule that maximizes the value of the products produced, while obeying engineering constraints. Major advancements in strategic mine planning have been accomplished over the past two decades allowing for the simultaneous stochastic optimization of a mining complex with stochastic mathematical programming. The simultaneous stochastic optimization framework aims to globally optimize the critical components in a mining complex including the extraction sequence, destination policy, blending, stockpiling, operating mode, capital expenditure, waste and tailings management, and transportation decisions while managing uncertainty (Montiel and Dimitrakopoulos 2015; 2017; 2018; Goodfellow and Dimitrakopoulos 2016; 2017; Levinson and Dimitrakopoulos 2020b; Paithankar et al. 2020; Saliba and Dimitrakopoulos 2020; LaRoche-Boisvert et al. 2021; Brika et al. 2023). The objective function is composed of two main parts. The first part aims to maximize the value of the products sold by accounting for the discounted revenues and costs that are accrued across each location in the mining complex. The second part manages technical risk by minimizing deviations from production targets using a set of recourse variables that account for uncertainty. This follows the pattern of a generalized two-stage stochastic programming formulation where the optimal decisions for the first-stage decisions are made based on the uncertain data available and the recourse variables measure the cost of deviations after uncertainty is revealed (Birge and Louveaux 2011). Optimizing the long-term production schedule provides an overarching plan for operations and allows a mining enterprise to understand future production and financial implications by modelling and forecasting the impact of future decisions.

Simultaneous stochastic optimization demonstrates the ability to significantly improve net present value and manage risk in long-term production schedules (de Carvalho and Dimitrakopoulos 2019; Kumar and Dimitrakopoulos 2019; Saliba and Dimitrakopoulos 2019; 2020; Levinson and Dimitrakopoulos 2020b; Paithankar et al. 2020; LaRoche-Boisvert et al. 2021; Morales and Dimitrakopoulos 2021; Brika et al. 2023) by moving away from conventional approaches that predefine the destination of a block based on the economic value of a block and ignore uncertainty (Dagdelen 2001; Caccetta 2007; Hustrulid et al. 2013). Instead, the optimization framework jointly determines the extraction sequence at each mine and the destination of materials based on the configuration of the mining complex and the uncertain quality and quantity of the materials in the ground. Material transformations and complex modeling related to geometallurgical behaviour are incorporated to account for the non-linearity related to the stockpiling, blending and the recovery of materials within the optimization framework (Goodfellow and Dimitrakopoulos 2016). The developments leading to the simultaneous stochastic optimization of mining complexes will be discussed through the remainder of this section along with the reasons for incorporating supply uncertainty through the use of stochastic orebody simulations.

Traditionally optimization approaches use a single estimated (average-type) orebody model to represent the material supply. The estimated model is assumed fixed and engineering and financial studies are completed ignoring the uncertainty and local variability of the material attributes in the mineral deposits. Studies show that uncertainty and local variability related to the material supply is the largest source of technical risk for mining complexes (Ravenscroft 1992; Dowd 1994; 1997; Dimitrakopoulos et al. 2002). Therefore, managing supply uncertainty when optimizing production schedules is of significant importance. Estimated orebody models, obtained by methods such as kriging (Isaaks and Srivastava 1989; Goovaerts 1997; Rossi and Deutsch 2014) smooth high and low grades by estimating the grade at points lying between known sample locations. This misrepresents the proportions of materials by smoothing the high and low end of the grade distribution. The effects of smoothing mislead the optimization frameworks as it cannot account for the direct impacts of supply uncertainty and local variations of the material grades. As a consequence, the production schedules obtained are unlikely to achieve forecasts. Stochastic optimization frameworks overcome this weakness by integrating uncertainty with a set of equally probable stochastic orebody simulations. Accounting for uncertainty and local variability using simulations that reproduce the statistics of the exploration sample data and other data sources are critical for managing risk related to supply uncertainty when optimizing mining complexes.

In the late 1990s, Newmont built a mine planning and process model to take advantage of several synergies in their Nevada mining complex (Hoerger et al. 1999a; Hoerger et al. 1999b). The mining complex contained several mines, stockpiles, and processing facilities with more then 90 metallurgical ore types and 60 different processing options. The proposed development was a mixed-integer linear programming solution for simultaneously optimizing multiple pits with several processing destinations in order to maximize net present value (Urbaez and Dagdelen 1999). The mining decisions included the timing of pushbacks and stope development, while considering capital expenditures and process specific trade-offs of cost versus recovery. The formulation showcased the ability to capture value by optimizing several mines and processing facilities in a single combinatorial optimization problem that blends material to satisfy constraints, meet production targets and reduce costs. However, the method still suffered from several drawbacks. The optimization framework required a large number of simplifications to reduce the model size including aggregating blocks into homogenized parcels, fixing the extraction sequence, and eliminating strategic decisions ahead of time. These developments were the foundations to the simultaneous optimization of mining complexes and led to substantial improvements over traditional mine planning approaches that typically follow a stepwise procedure. However, an estimated orebody model was still used as input into the optimization framework leading to plans that are incapable of managing supply risk.

Conventional long-term optimization approaches use a single estimated orebody model as input into the optimization framework. Lane’s cut-off grade optimization is performed to determine the optimal cut-off grades (Lane 1964; Lane et al. 1984). An ultimate pit limit and pushback design is completed for each mine independently (Hustrulid et al. 2013). Then, the annual long-term production schedule is optimized restricted by the output of each previously made decision. The local optimization of these major production scheduling decisions can significantly impact the value of the production schedule, as each business decision is based on the predefined limits of the prior output. The primary reason for performing the optimization in steps in the past was to provide a computationally efficient method by dividing the solution approach into a series steps. However, major assumptions are made when optimizing the ultimate pit limit, pushbacks and cut-off grade decisions separately which can lead to lower valued production schedules.

Similar to the work completed in Nevada, Whittle and Whittle (2007) and Whittle (2018) also develop a framework for globally optimizing mining complexes, which includes multiple pits, underground mines, stockpiles, blending constraint and different processing routes. The global optimization framework groups together nested pit shells to satisfy a stripping ratio and prioritizes them based on the economic value given the estimated inputs for each mine independently (Whittle 2007). Again, mining blocks are aggregated. The affect of aggregation limits the flexibility of the schedule leading to sub-optimal solutions by grouping materials into panels and further subdividing panels into parcels of material with similar qualities, which are assumed homogeneous (Whittle and Whittle 2007). The algorithm then proposes a random feasible panel extraction sequence and with these fixed decisions the algorithm optimizes the downstream processing, blending, and sale of products to maximize net present value (Whittle 2018). It then locally optimizes the best solutions by modifying the start and end period of each bench and reoptimizes, searching for a globally optimized solution. When extracting from the panels, it is assumed that an equal portion of each of the parcels will be retrieved, which ignores the impacts of material selectivity. Throughput and grade recovery relations are integrated into the block model to improve the selection of destinations for different parcels during the optimization (Whittle and Whittle 2007). Although, this method breaks the problems into a series of steps and locally optimizes several parts within the global optimization framework. The proposed method can consider a large number of material types and processing destinations allowing it to optimize the long-term production schedule for large mining complexes. The major limitations of this approach are that it uses a predefined ultimate pit-limit and phase design and ignores supply uncertainty.

The previously discussed methods aim to globally optimize mining complexes using estimated orebody models. The limitations of using estimated orebody models are addressed through the use of stochastic orebody simulations. Early stochastic optimization approaches attempt to integrate supply uncertainty when optimizing a mine production schedule by using the expected value of a block and a probability of being above a predetermined cut-off grade using a linear programming model (Dimitrakopoulos & Ramazan, 2004). The probabilistic approach ensures blocks in more certain areas of the deposit are mined earlier and links to the effects of risk discounting, where blocks with higher risk are deferred to later periods. The probabilistic approach also includes constraints that generate mineable extraction sequences by using equipment accessibility and mobility constraints that smooth the production schedule. A case study completed at a multi-element nickel-laterite deposit shows lower risk of meeting ore production targets with the proposed model. A limitation of this approach is the probability of exceeding the cut-off does not account for the joint uncertainty of large groups of blocks simultaneously instead it considers each block independently using a predefined probability cut-off. However, the equipment accessibility and mobility constraints are highly advantageous for producing a feasible long-term production schedule.

An alternative to the probabilistic approach for considering the joint local supply uncertainty was developed by Godoy and Dimitrakopoulos (2004) using a framework based on simulated annealing (Kirkpatrick et al., 1983). A multi-step optimization uses stochastic orebody simulations to retrieve a schedule that minimizes deviations from ore and waste production targets. The process begins by creating a series of optimal mining schedules for each simulation of the orebody using best and worse case mining schedules. To find the stable solution domain, the common area of cumulative ore and waste graphs are found by optimizing a set of equally probable stochastic orebody simulations independently. Then, with the stable solution domain, a linear program is used to produce a schedule for ore production and waste removal to retrieve the optimal mining rates by maximizing the discounted cash flow. The mining rates are used to optimize a production schedule for each stochastic simulation independently. These schedules are technically feasible and maximize net present value for each stochastic simulation independently. Then, a stochastic production schedule is found by optimizing the schedule with simulated annealing. The algorithm fixes the schedule of blocks with a high probability of being in a single production period. Then, unfixed blocks are swapped between periods to minimize deviations production from targets for each mining sequence over all the periods. Although, the proposed method indirectly integrates uncertainty into the optimization process it can not consider synergies between components and is limited by only having the flexibility to change the scheduled period of uncertain blocks. In addition, the model is unable to account for grade blending or control the risk of deviating from production targets. It also suffers from computational requirements that are alleviated by using a pre-determined probability cut-off grade, which has a large effect on the outcome of the production schedule. The optimization approach is applied on a case study at the Fimiston open-pit in Western Australia and leads to a 28% increase in net present value.

Albor Consuegra and Dimitrakopoulos (2009) complete a sensitivity analysis with the method previously introduced showing that 10 stochastic orebody simulations provide a stable production schedule and that the size of the pit tends to be larger when compared to a conventional approach. In addition, the framework is tested in a copper deposit showing 26% increase in net present value in a deposit with low-variability emphasizing the advantage of using simulations of the material supply instead of a deterministic orebody model (Leite and Dimitrakopoulos 2007). The study also investigates the impact of a risk discount rate and finds that it does not drastically change the production schedule when comparing different rates. Whereas the probability threshold required for discriminating ore and waste can lead to very different schedules and stripping ratios, identifying a major drawback in the need to decide an acceptable probability cut-off.

Ramazan and Dimitrakopoulos (2004) introduce a two-stage stochastic programming approach that uses a set of conditionally simulated orebody realizations as input to address the limitation of the simulated annealing based framework by Godoy (2003). Stochastic programming is a form of mathematical programming that determines optimal decisions to problems that involve uncertain outcomes (Birge & Louveaux, 2010). The objective function maximizes net present value and minimizes the risk of deviating from production targets. Geological risk discounting is applied to the penalty costs for deviations to defer risk to later production periods. The first stage decisions optimize the mining extraction sequence, and the second stage recourse variables quantify the risk of deviating from production targets based on the uncertain outcomes quantified using a set of stochastic orebody simulations. The framework manages grade blending requirements and maximizes the expected net present value. The primary drawbacks of this early stochastic programming formulation are that it precomputes block economic values, the destination of ore material is predetermined and can only consider a single processor, there are also no stockpiles. Furthermore, the stochastic programming formulation is applied to a two-dimensional data set that is not representative of a real mining operation.

The two-stage stochastic programming model is extended to include stockpiles and the flexibility to add multiple processing streams by adding terms to the objective function (Ramazan and Dimitrakopoulos 2013). The stochastic programming model requires a set of stochastic orebody simulations to optimize the production schedule and the optimization formulation aims to control the magnitude of risk associated with the material supply by modifying production scheduling decisions. The limitation of this approach is that it can only consider a single mine and processor. Additionally, the approach pre-calculates the economic value of a block in each simulated realization.

Leite and Dimitrakopoulos (2014) test the previously introduced stochastic programming framework at a low variability copper deposit to see the impact of the scheduler on a different mineralization type. The resulting schedule led to a 29% increase in net present value when compared to a conventional optimization approach and decreased technical risk. Benndorf and Dimitrakopoulos (2013) also apply the stochastic programming approach at a multi-element deposit and incorporate mineability constraints like the work of Dimitrakopoulos and Ramazan (2004). The study shows that the stochastic model improves the blend of uncertain silica and alumina grades and reduces risk when compared to the outcome of a conventional optimization approach, obtained with Blasor BHPs mine planning tool (Stone et al. 2005).

Menabde et al. (2007) develop a mixed integer linear programming formulation that uses a set of stochastic orebody simulations as input and simultaneously determines the cut-off grade and extraction sequence decisions in a single framework. The paper showcases that simultaneously optimizing the cut-off grade with other production scheduling decisions can lead to substantial improvement in the expected net present value. A continuous decision variable is used to determine the fraction of a panel to be extracted with a particular cut-off grade in each period. The cut-off grade decision is made by discretizing the grades into a number of pre-determined bins and determining the cut-off grade decision that maximizes net present value. The main limitations of this approach are: (1) it groups blocks into panels and assumes that fractional extraction is possible, however, the proportions of material accessible over multiple time periods are likely not be equal throughout the panel; (2) the formulation does not control risk over time but only ensures the average ore and mine capacity is less then the maximum capacities; and (3) the use of bins is not extended to manage multi-element deposits. The mixed integer programming formulation does provide a technique for simultaneously optimizing cut-off grade and extraction sequence decisions in a single optimization formulation. Similar considerations can be integrated into the stochastic programming formulation that accounts for risk and optimizes blocks instead of aggregated panels.

Rimélé et al. (2018) further advance the stochastic programming model to include the possibility for in-pit dumping; a method for reducing mining costs and minimizing ex-pit waste production. The approach extends the deterministic framework developed at BHP where a sequence is determined for block aggregates where waste blocks can either go to a waste dump or dumped in-pit (Zuckerberg et al., 2007). The approach considers truck hours int the formulation to effectively define the mining rate based on the hauling distance to different destinations and a constraint is included for the capacity at the waste dump facility. A set of mining areas named strips are available for storing waste material and constraints are enforced to determine the amount of material that must be extracted from a strip before it can be used for storage. For each period, the storage area is defined by a maximum of one top and bottom strip. A strip is only available for storage if it is located between the top and bottom strip, and this delimits the in-pit dumping area. Lastly, there is a maximum storage capacity for each strip and the algorithm decides whether or not a strip is required to be used for storage. The resulting formulation is applied in an iron ore mine in Labrador, Canada and effectively stores tailings and waste inside the pit. In-pit dumping substantially reduces the costs of rehabilitation and decreases the waste dump footprint. The main weakness of this approach is it is designed for large low-dipping mineral deposits and with very specific in-pit dumping criteria. Additionally, the work uses an exact solving method with a very small example using 3,177 blocks, which is not sufficient for long-term strategic mine planning and may not be scalable if more periods and blocks are considered. More efficient solving approaches such as metaheuristics must be considered.

Mai et al. (2019) suggest a stochastic programming approach for optimizing a long-term production schedule that aggregates blocks into predefined cones instead of panels. The approach significantly reduces the number of integer decision variables in the stochastic programming framework allowing the problem to be solved in a shorter timeframe. A linear program combines mining blocks into cones based on the expected value derived from the estimated orebody model (Mai et al. 2018). Cones store the quantity of ore and waste tonnage in each stochastic simulation and the expected economic value. This framework is very similar to the approach seen in Ramazan and Dimitrakopoulos (2013) except that it aggregates blocks into cones. Cones utilize estimated economic values, which fail to account for uncertainty of material grades and potential for larger pits. The number of cones is predetermined along with the minimum size of the cone potentially leading to user defined parameters that can drastically impact the outcome of the optimization. The cones do not account for suitable mining widths. In addition, artifacts tend to appear in the schedule where the outlines of the pits are highly concentric. Although aggregation approaches can improve solution times, it deteriorates the value of the production forecasts by combining blocks into fixed shapes that must be mined together. Block based optimization approaches can be slower, however, additional value can be obtained by allowing for further flexibility of the mining sequence and maintaining an understanding of the selective mining unit and its behaviour within a mining complex. This is the major reason for exploring alternative solution methods, such as meta-, hyper and math-heuristics to achieve faster solution times and overcome the desire to reduce the problem size with pre-processing steps (Lamghari and Dimitrakopoulos 2012; 2020; Lamghari et al. 2022).

Montiel and Dimitrakopoulos (2013) present an extension to the multi-step optimization approach presented in Godoy and Dimitrakopoulos (2004) by accounting for multiple processing streams, stockpiles and products. The objective function minimizes deviations from productions targets over all scenarios, processes, and periods. Each process plant accepts certain material types with different costs and recoveries. The method is applied at Escondida Norte, Chile a deposit that contains several sulphide and oxide material types, which can be processed through three different treatments: a milling process, bioleaching plant, and an acid leaching plant. Stochastic orebody simulations account for uncertainty in the copper grades and the material types. The optimization approach modifies the production schedule to minimize deviations from blending and production targets, however, it does not simultaneously maximize the net present value and is limited to a sequence where only uncertain blocks can be interchanged between periods. In addition, it only assigns a destination to blocks that are invariable in material type to prevent incompatibilities when a block may a different material type in different realizations. This approach suffers from several limitations as it requires a good initial mining sequence, it does not explicitly maximize net present value but tries to meet pre-determined production targets and assumes perfect selectivity of material types across scenarios leading to optimistic forecasts. The method does provide a means for considering several material types and processing options in a single optimization framework with various recoveries, which is highly advantageous for optimizing the long-term strategic mine plan.

Montiel and Dimitrakopoulos (2015) propose a non-linear two-stage stochastic programming model to simultaneously optimize mining complexes. The model optimizes the extraction sequence for multiple mines along with processing and transportation schedules in a mining complex. The mathematical formulation considers supply uncertainty with a set of stochastic orebody simulations that account for both material and grade uncertainty. An initial solution is created using a conventional optimizer and simulated annealing perturbs three different decision neighbourhoods that modify the blocks mined in each period, the operating alternative, and transportation system. Different from the methods discussed up until now, the approach calculates the discounted profit for each period by maximizing the revenue obtained from the recovered products and minimizing the cost of transporting and processing materials at each of their respective locations. The method maximizes the value of the products produced and does not pre-calculate the economic value of the block overcoming a major limitation from the previously mentioned stochastic programming approaches, as this allows for the value of interactions between different components to be captured. Stockpiles are created for each material type that has a different metallurgical property. The material at the stockpiles is blended to give an average grade forming a non-linear constraint that is feasible due to the metaheuristic solving approach applied. Each operating alternative has its own capacity, unit cost, recovery, metallurgical property requirements and throughput specifications. The transportation systems are limited to being used for certain process and must be capable of transporting all the material output from a given destination. By introducing several new components in the mining complex, the problem becomes substantially larger to solve with numerous binary decision variables. Simulated annealing is applied to solve the stochastic programming model (Metropolis et al. 1953; Kirkpatrick et al. 1983). To avoid reaching local optima and allow for sufficient exploration of the solution domain, simulated annealing allows for the solution to deteriorate and diversifies the search neighbourhood by perturbing different decisions levels. The probability of accepting a solution that is unfavourable is governed by the both the temperature, which is iteratively decreased during the optimization, and the change in the objective function due to the selected perturbation. Higher temperatures result in a higher probability of accepting an unfavourable decision using the simulated annealing algorithm. A case study is completed at a mining complex with two open-pit mines and five different material types each having a different recovery at four processing facilities. Synergies are found between the operating modes and processing alternatives given the material from multiple pits and improve the ability to meet production targets and increase project value. A 5% increase in net present value was obtained due to the solvers ability to satisfy the two mill capacities and blending constraints when compared to a conventional optimization approach.

Montiel and Dimitrakopoulos (2018) apply the stochastic programming framework for mining complexes at Newmont's Twin Creeks operation which shows improved metal forecasts, increased cash flow and lowered risk of achieving production and blending targets leading to a 6% increase in net present value. The multi neighbourhood simulated annealing approach is applied to this case study with perturbations that modify the extraction sequence bench-by-bench, change the destination of blocks, and adapts the stockpile rehandling schedule. Although a major advancement, a limitation of this approach is that the perturbations require an initial solution provided by a conventional optimizer to reach a good solution. Perturbations could be improved to rapidly create an initial production schedule with larger changes and later smaller perturbations can be applied to improve the production schedule. Additionally, only block modifications are applied making it difficult to explore new regions; therefore, alternative diversification strategies should be considered.

Goodfellow and Dimitrakopoulos (2016; 2017) present a method for simultaneously optimizing a mining complex by considering blending, non-linear recoveries, multiple mines and processing facilities. A generalized two-stage stochastic optimization approach is developed where several combinations of metaheuristics are tested including particle swarm optimization (Kennedy and Eberhart 1995), differential evolution (Storn and Price 1997) and simulated annealing (Kirkpatrick et al. 1983). The formulation maximizes the value of the products sold, while simultaneously managing the risk of deviating from production targets in a single mathematical formulation. The shift from focusing on the economic value of a block to the value of the products sold allows the optimization to consider complex non-linear and non-additive geo-metallurgical behaviour in the processing stream and stockpiles. Therefore, the blending and mixing of materials can provide additional value, which is not possible when assuming blocks are processed independently and have predefined destinations. Blending materials from several sources to understand the behaviour of non-linear recovery curves and hardness impacts at the processing plants can be considered. The generalized stochastic mathematical programming formulation proposed is intractable to solve with conventional stochastic mathematical programming methods due to the non-linear functions that are considered when optimizing a mining complex and the proposed generalized formulation. Therefore, a simulated annealing framework is applied and compared to the performance of both the particle swarm and differential evolution optimization methods. Particle swarm optimization and differential evolution were found to be highly sensitive to the initial sequence and destination policy chosen and also are computationally expensive for determining the extraction sequence. The methods were instead applied to optimize the destination policies and downstream optimization decisions while, using simulating annealing for determining the extraction sequence. A modified simulated annealing algorithm was used for selecting the perturbation strategy to be applied. The perturbations include randomly changing the blocks mining period, the destination policy, and the proportion of material to be moved between downstream components. The modified simulated annealing algorithm maintains a cumulative probability distribution function for each perturbation and instead of using a single temperature for each neighbourhood, a single parameter is used which represent the probability of accepting a non-improving solution. The temperature variable is found by looking up the temperature value using the cumulative probability density function for each neighbourhood. The method is tested on a copper-gold mining complex where it is found that the simultaneous optimization of the mining complex leads to a 22.6% higher net present value when compared to a conventional optimization approach. In addition, the optimization with both simulated annealing with differential evolution and simulating annealing with particle swarm optimization led to improved results of 2.57% and 1.91%, respectively. These marginal improvements came at the cost of running it for 2.9 and 2.4 times longer than the modified simulated annealing alone.

Goodfellow and Dimitrakopoulos (2016) cluster blocks into groups with similar attributes. The group membership of a block may differ between simulated orebody models, however, materials within a group are sent to a single destination across all scenarios. K-means++ clustering algorithm is used to determine the cluster centroids (Arthur & Vassilvitskii, 2007). The number of clusters is a user-defined input and describes the flexibility of grouping materials. As the number of clusters increases the number of decisions variables approach a block-based destination policy. However, using a small number of clusters a destination policy can be constructed that substantially reduces the number of decision variables and leads to good solutions to the long-term production schedule, which can be used in practice and remain stable when different simulations are used. The advantage of this approach is that k-means++ can be applied to multivariate deposits to cluster blocks based on several attributes. The destinations of a block may change between scenarios depending on the simulated attributes; however, the clusters and groups remain fixed throughout the duration of the optimization representing the operational selectivity available at the time of mining.

Capital investment decisions are also critical components in strategic long-term planning. Capital investments in equipment and infrastructure govern the ability to transport, crush, and process raw material to generate a product in a mining complex. Goodfellow and Dimitrakopoulos (2015) develop a two-stage stochastic integer programming approach that directly includes capital investment decisions to increase and decrease capacities using the generalized approach discussed previously. The mathematical formulation includes the cost of undertaking a capital expenditure in the mathematical formulation and uses a unit-based decision variable to increase the mining capacity by investing in trucks and shovels. A decision variable defines the number of capital expenditures that are undertaken each year and the model accounts for appropriate lead times and equipment life. The simultaneous stochastic optimization framework outputs a schedule that provides an investment plan along with the extraction sequence and destination policy necessary for optimizing the mining complex. A case study is completed on a copper mining complex that results in a 5.7% increase in net present value when compared with a conventional approach. This approach improves the ability to optimize the mine production schedule as feasible investments can be considered while the scheduling of the open pit mine is taking place, leading to plan adaptations that may have not been considered otherwise. In addition, it provides an understanding of the optimal time to take on additional investments to increase the value of the mining complex.

Kumar and Dimitrakopoulos (2019) integrate geometallurgical, grade and material type uncertainty into the simultaneous stochastic optimization of a mining complex. The application simulates the semi-autogenous power index (SPI) and bond work index (BWI) of materials in the deposit and the hardness is derived from these properties, yielding 12 material types. Geometallurgical targets are created to maximize utilization of the processing facilities. The application of the simultaneous stochastic optimization approach is tested at a copper-gold mining complex considering new hard/soft ratio targets resulting in improvement at the processing facilities. This led to a 19.3% increase in net present value when compared to a conventional production schedule which is driven by the ability to incorporate material hardness into the simultaneous stochastic optimization framework.

Saliba and Dimitrakopoulos (2019) assess the impact of joint supply and commodity price uncertainty by integrating them into the developments of the generalized two-stage stochastic integer programming approach discussed previously. The simultaneous stochastic optimization is applied at a multi-pit, multi-processor mining complex with critical blending constraints. The case-study shows that incorporating market uncertainty into the simultaneous optimization of a mining complex leads to more material being processed when higher prices are more probable, whereas the schedule is slightly more conservative when there is a higher chance of a decreased price. The improvement of the P-50 (median) when comparing the schedule produced prior to considering price uncertainty and the schedule that incorporates price uncertainty produce a similar net present value however, the spread between the P-10 and P-90 increases drastically highlighting the technical risk of commodity price uncertainty. The limitation of this approach is there is no explicit method of managing the price uncertainty in the optimization framework, therefore, the schedule only adapts to the upward and downward trends of the price fluctuations. Saliba and Dimitrakopoulos (2020) present another application using the same formulation to account for tailings management. However, they include capital investment decisions. The case study considers a capital investment opportunity to expand the tailings management area and extend the mine life at an operating multi-pit gold mining complex. The case study simultaneously chooses the optimal time to expand the tailings facility and increase the operational life of the mine leading to a schedule that increases metal production and net present value by 14% and 4%, respectively.

Fu et al. (2019) develop a mixed integer programming model that determines the optimal production and waste-rock dumping schedule with an objective to maximize the net present value of an open-pit mine. The novel approach integrates haulage distance from each block to various destinations based on the distance between the pit exits and processing plant, waste dumps and stockpile entries. In addition, the model uses a dynamic cut-off grade concept and can stockpile both ore and waste material. For each destination it applies the dynamic cut-off grade by determining all possible economic values of a block. A block-based approach is applied to solve the optimization problem that requires pre-calculating all costs to determine the appropriate destination. Thus, the method is not capable of accounting for complex blending and non-linear recovery curves. Although the destination is chosen during the optimization having more then one intermediate destination results in computational challenges and an extension of the formulation. The proposed model allows for partial block mining to enable the use with parcels and panels, which has its limitations that have been discussed. In addition, the stockpiling framework discretizes the stockpile bins into bins, where each bin accepts materials with a given interval. They assume the average grade retrieved from the bin will be exactly the average of that bin which can lead to problems for instance sending all material to a stockpile bin with bounds between 1 and 1.5 and an average grade of 1.25 could be used to send all material between 1 and 1.25 and upgrade the value of the material. This could lead to misguided forecasts with higher then expected metal recovery due to artificially created metal. The model accounts for the placement of non-acid generating and potentially acid generating waste materials, directly incorporating waste management and the appropriate locations to send material during each time period. This adds significant value as it can reduce the effects of acid rock drainage a common environmental problem by selectively placing and encapsulation potentially acid-generating waste rock (Fu et al. 2015). However, due to the many precedence constraints and the use of exact solving methods the case study is extremely small, only considering 1,688 mining blocks and a four-year time horizon which is not sufficient for a large-operating mining complex. Lastly, it fails to account for uncertainty.

Levinson and Dimitrakopoulos (2020b) simultaneously optimize the long-term production schedule in an operating gold mining complex while considering waste management. The simultaneous stochastic optimization framework controls the production of uncertain potentially acid-generating waste material by simultaneously optimizing the extraction sequence, cut-off grade, and stockpiling decisions, and actively managing uncertainty. The simultaneous approach reduced the cost of reclamation and better adhered to environmental constraints. The resulting schedule shifts the focus to high value areas and mines less material reducing the total amount of waste material processed. The processing stream is more effectively utilized leading to a 6% increase in net present value. This contribution towards the net present value does not account for the likely possibility of reducing the fleet size and incurring lower reclamation costs due to a reduced waste dump footprint. Each of these could be considered as beneficial results of simultaneously optimizing the production schedule with waste management.

Paithankar et al. (2020) develop a mathematical formulation to optimize the extraction and destination policy of a mining complex under supply uncertainty. The stochastic programming approach uses a genetic algorithm for the solution approach (Dengiz et al. 1997; Mitchell 1998; Goldberg 2013) and integrates Lane’s method for cut-off grade policy to determine the destination policy (Lane, 1964). The ultimate pit-limit under geological uncertainty is solved using a graph structure and calculating the economic value of the block in each scenario for each processing destination (Paithankar & Chatterjee, 2019). Block are classified as ore or waste based their economic value and hybrid maximum flow and genetic algorithm is used to find the ultimate pit where the genetic algorithm optimizes the weights of the arcs in the graph like structure. Given the mining extraction sequence a set of binary decision variables define the destination of a block in each scenario. Lane’s cut-of grade is calculated and considers the grade-tonnage curve finding the cut-off that maximizes the discounted cash flow while, satisfying mining process and refining capacities. To determine this, they use the quantity of ore to be mined, processed, and refined averages are taken across scenarios to utilize Lane’s relationships. Therefore, the cut-off grade is optimized to satisfy the average of the scenarios and does not directly maximize the profitability under uncertainty. In addition, the method is optimistic as it assumes the perfect execution of block selectivity based on the exact cut-off grade decisions, which are unlikely to be realized operationally and only consider the average across scenarios. The multistage approach is solved in a series of stages to reduce the problem size utilizing a global objective and still only considers a small problem with 12 periods and 44k mining blocks. Lastly, mining width is not considered leading to impractical mining production schedules.

Del Castillo and Dimitrakopoulos (2019) expand the two-stochastic integer programming framework for simultaneously optimizing a mining complex to develop an adaptive optimization approach that considers investment decisions but, allows the production schedule to branch depending on the probability of undertaking the investment. Similar to a multi-stage stochastic programming approach a branching mechanism is described that chooses the opportune time to invest and provides a number of investment choices along with their impact to the production schedule. The branching mechanism is based on the probability of undertaking an investment in different groups of simulated orebody realizations and when the probabilities are nearly equal the schedule is allowed to branch and adapt the schedule based on the investment undertaken in each branch. This is accomplished by introducing non-anticaptivity constraints that ensure that each branch the production schedule is the same for all scenarios within the branch. In addition, a representativity measure is used to reduce the dimensionality of branching and ensure a feasible production schedule is created, overcoming limitations of assuming there is a schedule per a simulated orebody realization (Boland et al. 2008). The framework is applied on a multi-mine copper mining complex that considers not only truck and shovel purchases but also opportunities to add additional crushers to increase the processing capacity. The adaptive framework leads to a $170M increase in net present value compared to the non-branching stochastic integer programming approach.

Levinson and Dimitrakopoulos (2020a) extend this framework to consider opportunities to expand the tailings facility, increase the capacity at an autoclave and allow for additional acid consumption. The resulting schedule branches on the autoclave investment leading to a 6.4% or 27.5% increase in net present value for the branch that takes the investment versus the case without taking the investment, respectively. The branching framework allows the schedule to be adapted to different possible outcomes leading to a higher net present value and a rapid method for evaluating the placement of infrastructure given different potential outcomes. Allowing for adaptations is extremely useful for mine planning applications over long-time horizons as more information will likely become available during mining and this provides more flexibility in the given plan. The method, however, suffers from its ability to find optimal solutions with large numbers of investment alternatives; therefore, smart perturbation techniques should be considered that allow time for the extraction sequence to adapt to the larger capacity changes.

Dimitrakopoulos and Jewbali (2013) introduce a multistage production scheduling framework that addresses compliance between short and long-term production schedules and the influence of short-scale information. Future grade control data is simulated with grade control and exploration data from previously mined areas of the deposit. Then the existing stochastic simulations are updated with successive residuals (Jewbali and Dimitrakopoulos 2011). Long-term production schedules are typically optimized with a set of stochastic orebody simulations. The simulations only account for exploration data, as planning occurs prior to collecting grade control information. Integrating grade control information provides additional information leading to better predictions of the geological characteristics of the material to be mined. Successive residuals update the stochastic orebody simulations with future grade-control data. Then, by optimizing the production schedule, a long-term production schedule that accounts for future information demonstrates improved better compliance with the short-term production schedule and reconciles closer than the schedule optimized with the original stochastic orebody simulations. A stochastic programming framework, similar to Ramazan and Dimitrakopoulos (2013), penalizes the objective for deviations that impact short-term production scheduling decisions, which relates to the goal of optimizing the short-term production schedule within the long-term scheduling framework. Accounting for future grade control information led to an additional 3.6M tonnes of ore and 2.6 M g of metal produced, which reconciled closer than the original set of stochastic orebody simulations. Therefore, accounting for additional grade-control information led to an informed production schedule that effectively managed risk with stochastic simulations that account for future grade control data. The principle of integrating the short-term production schedule and accounting for new-incoming information, in this case as grade control, provides an advantageous way to adapt the long-term production schedule to improve the ability to meet production targets, maximize net present value and mitigate risk. Future grade control data is not the only type of information available, and the benefits of addition infill drilling data could also improve the ability to plan the long-term production schedule.

Peattie and Dimitrakopoulos (2013) present a simulation-approach for recoverable reserve forecasting that incorporates future grade control data by adding simulated correlated random sampling errors to the original simulations based on information from previously mined areas. A case study is completed at the Morila gold deposit in Mali to assess production schedules, which shows that the original simulations with incorporated grade control data closely match the actual grade-tonnage proportions from grade control. Although, methods such as this provides an accurate risk analysis of the expected reserves and production schedule it does not consider a method for minimizing the deviations from forecasts. The multi-step approach discussed here uses short-scale deposit information in the form of grade control simulations alongside the set of simulated orebody realizations as input into a two-stage stochastic programming formulation. The two-levels of information provide a way to incorporate grade control information into a model to manage the risk of local variability and local classification of materials, ore and waste, and the areas to be scheduled over the life-of-mine. However, the scheduling benefits were not addressed to see how this method could impact production scheduled decisions.

Kizilkale and Dimitrakopoulos (2014) develop an optimal control framework that considers financial uncertainty while optimizing the mining production rates for multiple mines in a mining complex. The dynamic program determines the optimal production rate for each year at each mine based on an uncertain price process. A distributed policy iteration method is run and is shown to converge to a unique Nash equilibrium with a policy that provides a suitable mining rate for each of the mines. The target extraction rate function is approximated with a parametric approach that considers stochastic prices and the mine extractions rates. The key component of this work is the mines are not independent and share constraints and stockpile capacities. The method optimizes the mines independently and with a finite number of iterations converges to a solution that considers the interaction between all the mines simultaneously. A limitation of this approach is that it does not account for supply uncertainty and only provides a mining rate not a mineable production schedule. In addition, the downstream decisions in the optimization framework are not considered. An interesting finding in this work is that under financial uncertainty the interactions between different mines do not behave proportionally to each other when optimized together whereas when optimized independently they are less sensitive to price uncertainty highlighting the importance of simultaneous stochastic optimization.

Zhang and Dimitrakopoulos (2017) develop a dynamic material value-based decomposition method that optimizes the mine production schedule and the downstream material flow plan to maximize the expected net present value. This method considers a mining complex with multiple mines. The aim of this work is to allow the optimization of the mine production schedule for each mine and the material flow plan to be synchronized through iteration. The model integrates both geological and market uncertainty. The mine production schedule optimizes the extraction sequence while considering supply uncertainty. Each block destination is assumed to be scenario dependent assuming perfect selectivity of blocks. Whereas the material flow problem accounts for the movement of material through the different components and is solved separately accounting for price uncertainty. The two optimization problems are then synchronized though iteration.

Zhang and Dimitrakopoulos (2018) also design a two-stage stochastic programming formulation to optimize a mining complex that considers both geological and market uncertainty. A heuristic is developed to manage non-linearities in the model including a throughput and grade dependent recovery. In addition, a method for evaluating and creating a forward contract is proposed to manage risk. Forward contracts require a mining enterprise to agree upon conditions for a particular unit price and quantity over a certain duration. The company is obligated to fulfill the contract. The contract is an alternative method to manage risk of market price uncertainty, guaranteeing a certain price for the materials to be produced. The model considers both spot market and current contracts in a single optimization problem. When a new contract is requested, it is added to the model and tested to see if it is worth undertaking by reoptimizing the mine plan. The upper bound and lower bound of throughput and material head grade at the processor are used in the heuristic approach. In each iteration of the optimization the method moves the lower bound and upper bound for each period and simulation toward the optimal throughput and product grade obtained. A parameter controls the convergence rate of the solution. Lastly, the approach considers whether it worth taking on a future contract based on set of simple pre-defined conditions. This is a notable addition to the simultaneous stochastic optimization of a mining complex as future contracts are effective method of managing risk in an uncertain environment. In addition, the ability to incorporate complex non-linear relationships into the optimization model using an efficient heuristic is an important advancement. A limitation of this approach is that the model needs to be run many times to determine optimal contract price and quantity, while considering market price uncertainty.

The simultaneous stochastic optimization paradigm provides a massive improvement on past work in its ability to optimize an entire mining complex in a single mathematical formulation. Many examples are shown highlighting the importance of integrating various components to determine the long-term production schedule. The results seen drastically improve the ability to meet production targets by incorporating and managing technical risk, while enhancing value by finding synergies between various components in the mining complex. This has been shown in examples that consider highly complex blending requirements, non-linear recovery curves, strategic stockpiling and waste management, capital investment considerations, and efficient methods for optimizing the cut-off grade for both multi-element and single element mineral deposits. Although, these methods have overcome various hurdles in optimizing a mining complex the simultaneous stochastic optimization framework must be developed further to consider a diverse number of waste management and environmental considerations. In addition, there are opportunities to consider the effects of material hardness and other geometallurgical properties on these processes. Preconcentration methods can also provide efficient alternatives that can minimize energy usage and help eliminate hazardous waste, which will be described in the subsequent section. In addition, progressive reclamation ensures that an effective mine closure strategy is considered.

1.2.2 Integrating preconcentration aspects into a mining complex

Mining enterprises persistently aim to integrate new and existing technologies to improve their ability to profitably deliver products to their customers. Environment and sustainability are at the forefront of their social agenda in a movement to improve current best-practice and ensure safe and viable long-term production. Several aspects that require consideration for sustainably producing valuable products include minimizing waste and tailings production, water and energy consumption, reagent and other harmful contaminant use, and carbon emissions. Preconcentration technologies include sorting, sensor-based bulk sorting and particle sorting can provide benefits by eliminating waste at a coarser scale then with a traditional processing facility. The impact of sorting and waste rejection technologies on production schedules has only recently been reconsidered and primarily focuses on small field studies and long-term production planning. The key developments in preconcentration technologies are discussed in this section.

Recent research in commercial preconcentration technologies aim to identify key material attributes that can be used to discriminate ore from waste (Espejel et al. 2017). For instance, the natural deportment of high-grade materials into specific size fraction after breakage can be utilized to target certain size fractions with higher metal concentrations for processing. Differential blasting has potential application when used to change the size distribution of ore and waste material by adjusting blasting parameters to induce deportment of high-grade materials (Salmi and Sellers 2021; Salmi et al. 2022). Target size fractions can be sorted and separated after blasting to prevent dilution and ore loss by exploiting the difference in size fractions within the ore and waste. Bulk sorting techniques use sensors in combination with methods to divert materials based on their predicted composition. X-ray fluorescence (XRF), laser induced breakdown spectroscopy (LIBS), magnetic resonance, electromagnetic, prompt gamma neutron activation analysis (PGNAA) and other methods have provided effective ways to detect physical and chemical properties of different materials (Hilscher et al. 2017; Nadolski et al. 2018). These engineering solutions provide opportunities to upgrade the material entering the processing facility by removing materials with low or no metal content prior to or during processing. This family of methods are referred to herein as preconcentration techniques and can lead to large environmental benefits including reduced energy and water use at the processing facilities along with decreased tailing production due to the early detection and rejection of waste in processing feeds. Although it is not relevant for all orebodies, certain ore types demonstrate large benefits using some or all of the proposed methods. Therefore, when applying these techniques, the natural material property can be exploited by separating the predicted low-grade material from the higher-grade material destined for processing to reduce costs and increase the input feed grade. Mineral reserves can also be increased by applying these techniques to upgrade a portion of the waste material, making it economical ore material after undergoing preconcentration. The main drawback with these technologies though is integrating them into the production planning process to assess and identify improvements in the short- and long-term production schedules linked to preconcentration decisions. This has received limited attention. This process requires an understanding of the uncertain material behaviour and the mechanisms required to eliminate waste throughout the mining complex using preconcentration technologies.

Burns and Grimes (1986) was one of the first applications of preconcentration technologies reported and was implemented at Bougainville Copper’s Panguna orebody. After blasting, material near the fracture planes in this deposit were found to have the copper and gold highly concentrated in the fine size fractions. In order to upgrade material and reduce costs, a preconcentration screening facility was installed with high throughput using banana type screens to annually preconcentrate 35 million tonnes of material near the cut-off grade. As a result, the feed grade to the processing facilities increases which extended the resource and allowed for lower grade materials to be mined.

Bowman and Bearman (2014) discuss the application of coarse waste rejection technologies that exploit the heterogeneity of material grades based on size. The paper discusses the implementation of a project completed at Newcrest Mining across there operations to reject coarse waste materials. The key environmental concerns identified include the significant energy consumption required for crushing and grinding ore materials prior to beneficiation and concentration techniques. This can be alleviated by removing waste materials intermixed with ore prior to processing based on some criterion. The second key consideration was determining cost effective separation methods that are not ignored due to the increase in handling costs of removing waste. One of the more cost-effective methods that can be used to separate material at a coarser scale is screening facilities for separating coarse and fine materials. The coarse separation study was completed at the Telfer gold deposit where it was found that approximately 75% of gold was hosted in 50% of the rock mass. Therefore, it was possible to reject 50% of the rock mass and recover 75% of the metal. This was based on screening and assaying the results from different size fractions.

More recent developments have focused on methods that model the expected behaviour for a given preconcentration technique. Carrasco et al. (2016a) define a coarse liberation model for early coarse rejection of waste material based on size. The coarse liberations model is used to measure the benefits of separating materials where metals tend to concentrate into specific size fractions after breakage. The magnitude of this behaviour is modelled through a response rank parameter (RR). The RR describes the relationship between the upgraded feed grade and the cumulative undersize weight by mass passing a screen. The geometallurgical parameter is defined by a series of laboratory tests on drill core and bulk samples. Using the RR and corresponding undersize mass, a response factor (RF) or upgrade factor can be calculated. Screen efficiencies impact the results during testing in an operating setting and should be compared with laboratory results. The RR model fitting is obtained through the least squares method. Carrasco et al. (2016a) complete a series of field tests for screening by natural deportment primarily looking into screening efficiency for separating materials of various particle size.

Preconcentration decisions have been shown to be quite advantageous and can be modelled and integrated into stochastic orebody models with geostatistical techniques. Moreover, these components must be integrated into the stochastic mathematical programming framework used for mining complex to accurately define the production schedule and understand the impacts of supply uncertainty. Preconcentration facilities transform the material properties depending on their material type, grade, and response rank likely leading to large differences in the resulting extraction sequence and destination policy.

Fathollahzadeh et al. (2020) present a mixed integer programming approach to generate a long-term production schedule that considers three preconcentration alternatives. In this study, fine under-sized material is considered high grade and is sent to the processing plant, or the leach pad and coarse oversized material is considered lower grade and can be sent to the heap leach or waste dump. Given the mining blocks attributes, a response rank is determined for the each preconcentration technique and the technique that maximizes the blocks profitability is chosen as the destination. Preconcentration decisions are included in the mathematical formulation used to optimize the long-term mine plan. Therefore, a number of data pre-processing rules are completed to solve the problem in a reasonable amount of time. The model also integrates constraints on steel-consumption in the ball mils and power usage. This mixed integer programming framework optimizes several components of a mining complex jointly; however, it only considers a single open-pit mine and precomputes the value of a block at each destination failing to allow for the interconnectivity between components. Several different mass recovery options are considered based on the material properties along with five different screening alternatives. The method assumes 307 grade bins in stockpiles to maintain the linearity of the problem, which is impractical in an operational setting. Furthermore, the largest instance of the problem solved was 20,142 blocks, which is small for solving a long-term production schedule in a large-scale mining complex. Computation results showed as the problem was scaled it failed to be solved. Lastly, the method fails to account for material uncertainty. To solve large scale mining complexes, the impact on the entire mining complex must be assessed and optimized to capitalize on the full potential of the available methods and better adhere to environment and sustainability guidelines. Metaheuristic techniques should be explored to integrate further components into the optimization on real-world mining complexes.

In addition to the developments of preferential grade by response using screening by natural deportment, opportunities exist to optimize blasting parameters to find a balance between fragmentation and operating costs to improve milling performance (Del Castillo 2018). Similar opportunities that modify the blast design to induce deportment of desirable size fractions with differential blasting show potential for major improvements (La Rosa 2017). One study shows that using screening and blasting techniques higher throughputs can be achieved if a smaller feed size fraction is used (Carrasco et al. 2017). Thus, improving the ability to separate high grade mass fractions from those of lower grade and separating low-valued mass fractions can potentially lead to a higher net present value and lower energy usage. Particle size analysis and ore sensing technologies can be used to improve the decision making process and further optimize mill throughput, while reducing excessive waste production and energy usage by considering laser induced breakdown spectroscopy (LIBS) (Bolger 2000) or mid-infrared and quantum cascade laser technologies (Parrot et al. 2020). These sensor technologies can be used to determine materials elemental composition and correlate it to the material type and grade to provide new informed decisions when short-term production scheduling. Adaptive frameworks will be required to quickly reoptimize short-term production schedules as new information becomes available to improve the decision-making process. These frameworks will help align the objectives of the long-term production schedule.

Determining the benefits of preconcentration techniques requires a comprehensive study of the material properties to understand the behaviour of material in the mineral deposit and the expected response factor. For example, hardness, grade, and particle size distribution based on the blasting and comminution parameters. All of which are highly uncertain in nature and are calculated by estimated models using a set of sparse samples. This provides an additional layer of uncertainty that relates to material type, material grades and operational aspects that are dependent on blasting, crushing, and screen efficiency. These uncertain aspects must be managed to minimize the effects of material misclassification and successfully preconcentrate materials. In terms of a mining complex, destination policies and stockpiling strategies must consider these properties. This can be accomplished by focusing on the value of the products sold and accounting for more complex interactions in the mining complex with different preconcentration approaches. Preconcentration technologies require a critical understanding of these mineral deposits and rely on unique characteristics which can not always be generalized across different types of mineral deposits. This information is often site specific.

An important aspect of strategic long-term mine planning is being able to take the production schedule designed and then schedule on shorter time horizons for instance weekly or monthly accounting for budgeting constraints and fulfilling the requirements of the forecasted production outlined in the strategic mine plan. This should be completed in a way to maintain project value but, also ensure operational constraints are satisfied and the appropriate use of different operating alternatives, such as preconcentration techniques, are effectively used given the available material in each production period. Material extracted from the mineral deposits is heterogeneous in nature which makes it difficult to manage the quantity of ore and waste requirements for satisfying processing during shorter time horizons. The proportions of material are not evenly spread through the mining areas and access is highly dependent on the short-term production schedule. Therefore, preconcentration techniques that separate materials by considering the material response rank in the mineral deposit can become even more beneficial at shorter timeframes by adding a layer of flexibility. Different metrics can be used on a short-term basis to maximize value and ensure operational aspects allow the production schedule to obtain long-term production targets. A method for integrating preconcentration into short-term production planning is developed in Chapter 3. Then, Chapter 4 discusses the opportunity to link short-term production scheduling within the long-term production schedule as method for delivering reliable forecasts that account for short-term impacts.

The remainder of this literature review will consider methods that optimize the short-term production schedule and apply reinforcement learning to rapidly produce the short-term production schedule. In addition, waste management, reclamation and recent works in modelling geological uncertainty using high-order statistics along with updating frameworks will be discussed.

1.2.3 Simultaneous stochastic optimization: short-term applications

Short-term simultaneous stochastic optimization is a framework that has not been explored to the same extent as its long-term counterpart. The concept of simultaneously optimizing a mining complex is maintained, while accounting for geological uncertainty and several other intricacies that are required to ensure feasible and operational mine plans. The breakdown of mine planning is typically broken into three types: strategic, tactical and operational. The applications seen up to this point aim to determine the strategic mine plan that involves simultaneously determining the long-term yearly extraction sequence, blending, processing, stockpiling, preconcentration, waste management, and capital investment decisions in a single mathematical formulation. The objective of the optimization is to maximize the net present value of a mining complex, while managing technical risk of the annual production schedule over the life of the operation. The strategic mine plan determines the physical constraints and limitations for planning at the tactical and operational level such as the areas to be mined, available equipment, blending and stockpiling strategies that maximize the profitability of an operation. In tactical planning, the focus is to execute the strategic mine plan considering a shorter timeframe which, tends to be days, weeks, months, or quarters. Tactical plans are designed to achieve operational targets including the quality and quantity of material to be processed and ensure fulfillment of the strategic mine plans. Available commercial solvers that exist typically use a mathematical programming approach to minimize costs and meet production targets. However, these approaches tend to forfeit value by failing to take advantage of new incoming information and ignoring the impact of geological uncertainty. At the operational level, a more detailed plan is made that includes fleet management, grade control and other daily operational constraints. For example, ensuring the processing facility is efficiently utilized, materials are blended to specification, waste dumps are effectively filled, and drilling/blasting is on schedule to maintain blasted inventories and ensure material is available for extraction. Short-term commercial planning packages have moved towards Gantt like activity scheduling that is not optimized and fails to account for possible synergies by solely relying on previously made decisions from the long-term production schedule. The majority of the work to follow will focus on the scheduling of a mining complex at the tactical planning level, however, an ongoing effort is being made to move towards real-time operational decision making that considers the mine production schedule and new incoming information coming from each component of the mining complex, along with fleet management. The conventional frameworks will be described to understand the key components of a mining complex to consider during short-term mine planning in addition to new innovative planning approaches that account for equipment and geological uncertainty with value driven objective functions. Short-term mine planning relates aspects from both the tactical and operational level where monthly or annual schedules are decomposed into shorter timescales to provide operational guidance and ensure compliance.

Supply uncertainty still remains a large risk to short-term production schedules and should be integrated into the production scheduling process to manage technical risk of meeting production targets. Although, more detailed information is collected through grade-control the material uncertainty is still widespread and must be managed accordingly. Shishvan and Benndorf (2016) further highlight the impact of geological uncertainty in a coal mine based on a short-term production schedule by completing a risk analysis using a process simulation-based approach. The risk analysis shows that estimated orebody models forecast 10% less coal than the average of the

20 orebody realizations used as input in the process simulation approach. The simulated orebody realizations are very similar when compared to the exhaustive model of coal quality. Once again, the estimated model is found to ignore local variability and uncertainty of the material qualities. Whereas the simulated orebody realizations are able to better reproduce the statistics of the data and the exhaustive model. The coal product delivered to the customer does not align with the predictions based on the estimated model leading to poor quality control. This emphasizes the need to account for uncertainty even in the short-term to ensure quality requirements have a high probability of being met by managing the inherent risk associated with geological uncertainty. In addition, Smith and Dimitrakopoulos (1999) apply mixed integer programming to optimize the short-term production schedule. The method showcases the influence of geological uncertainty by testing a conventional schedule with several different simulations. However, the optimization formulation fails to manage geological uncertainty leading to sub-optimal mine plans. The approach suggests a need to develop a stochastic programming framework.

Mann and Wilke (1992) incorporate grade control into short-term production scheduling to improve blending at the processing facilities and keep the limits of material within acceptable levels to prevent the production of hazardous by-products. Integrating grade control into short-term planning adds several layers of additional complexity due to the cumbersome accessibility constraints that relate to block precedence and shovel allocation decisions. The work highlights a trial-and-error approach that aligns with short-term planning software strategies that are commercially available. The mine planner iterates through different sequences essentially creating a short-term production schedule by defining the mining rate for each piece of mining equipment (Smith 1998). Using this approach, it is very difficult to create a schedule that is optimal or even near optimal without quantifying the impact of all possible production schedules. A linear programming model is proposed in an attempt to optimize the short-term production schedule to ensure daily ore production is meet and that quality requirements are as close to the strategic mine plan as possible. The main drawback of this early approach is it locally optimizes several aspects of the mine plan on a period-by-period basis, failing to model the scheduling and blending model requirements in a single formulation due to computational limitations. Therefore, they suggest an iterative approach where each period is solved independently and only the blending is optimized using the linear program, whereas shovel movements and scheduling are decided separately using a CAD like canvas. The method does highlight critical aspects that must be considered in the short-term production scheduling process, which include shovel movements, accessibility constraints and waste removal.

Several other early models for managing portions of the short-term production schedule are described. Wilke and Reimer (1977) use a linear programming approach to minimize deviation from a long-term production plan, while maximizing utilization of equipment and minimizing shovel movements. The linear program aims to meet the blending requirement of hardness, iron, magnetite and silica grades within an acceptable range. Chanda (1990) present a goal programming approach for the crushed stone industry that prioritizes each objective; however, the model is simply a supply and demand problem that does not consider scheduling the extraction sequence. Chanda and Dagdelen (1995) consider blending optimization using a goal programming approach combined with interactive graphics that allows the mine planner to change the schedule graphically and then run the blending model. Fytas et al. (1987) demonstrate a linear programming model that groups volumes to be scheduled based on mining areas. Each area is assigned a priority factor based on its accessibility and materials available for blending the products. The linear programming model requires block aggregation and allows for partial block mining, which does not account for the impact of selecting an appropriate selective mining unit. Smith (1998) use mathematical integer programming to optimize over multiple periods with access constraints with AMPL.

A major challenge when optimizing short-term production schedules is finding an optimal way to satisfy blending requirements on a daily, weekly, monthly, or quarterly basis as materials are heterogeneous. Therefore, blending constraints may be managed in the long-term yearly schedule, however, due to accessibility and equipment limitations it is not always possible to meet these requirements at shorter time scales. Kumral and Dowd (2002) further emphasize the importance of grade control and blending material to specific criteria to satisfy the short-term production schedule for industrial metals. The multistage process leverages a linear programming approach that identifies blocks that can be blended without violating grade requirements. Then, Lerchs-Grossman algorithm is applied to the blocks found in the linear program to determine the ultimate pit. Lastly, all the blocks within the ultimate pit limit are scheduled first using a Lagrangean parametrisation to retrieve an initial sub-optimal solution and then iteratively improve the schedule using a multi-objective simulated annealing framework. The framework aims to minimize deviations from blending targets, ore production targets, grade fluctuations and maximize the mine life. The method requires a series of simplifications including separating the problem into various stages and including only blocks that ensure blending constraints are satisfied. In addition, the simulated characteristics of each block are averaged to create the input model, failing to consider the impact of local uncertainty and variability of the simulated characteristics.

Eivazy and Askari-Nasab (2012) propose a mixed integer linear programming model which minimizes the cost of mining, processing, hauling, rehandling and rehabilitating by generating a production schedule that complies with a long-term production schedule. Buffers, directional mining, ramp access, stockpiles, and waste dumps are also included in the formulation. A monthly production schedule is created for an operating year at an iron ore mine. The model determines the extraction sequence and allows for fractional block extraction. The destination of the mined materials are based on material type, processing requirements, material grade and haulage cost. Lastly, stockpiles are considered. There remain several drawbacks to this approach. First, it groups blocks into aggregates and allows them to be sent to several destinations in an attempt to reduce binary decision variables and increase the flexibility of meeting blending constraints. Aggregating blocks and allowing for fractional block mining is advantageous for satisfying production targets and speeding up the optimization process. The challenge is operationally sending a portion of a block with an assumed homogeneous grade to two separate locations can be misleading. Defined cut-offs or destinations should be established to provide acceptable production guidance. The aggregation of blocks also further smooths the grade of blocks within an aggregate leading to a decreased understanding of the material behaviour and when reconstructed using the original block sizes can potentially lead to an infeasible mine plan (Mousavi et al. 2016b). Another weakness of this approach is that a constraint is defined to force all blocks to be mined during the scheduling time horizon, which can lead to infeasible solutions if operational constraints restrict production. Nevertheless, the model does propose a method for integrating rehabilitation costs based on waste production. It assumes a rehabilitation cost for each tonne of waste placed, which is limited as it does not account for actual reclamation activities undertaken during the short-term planning horizon and its interaction with equipment and construction requirements.

Upadhyay and Askari-Nasab (2018) demonstrate a simulation and optimization approach for short-term production scheduling of open pit mines. Operational uncertainty is considered by applying a discrete event simulation model of the mining operations, emphasizing the need to account for changes in the production environment and subsequently adapting the plan to minimize deviations from the original plan based on operational risk. The approach simulates the material handling activities to generate short term plans within the constraints of the long-term strategic mine plan. An optimization tool chooses shovel locations based on a strategic schedule to link operations directly with the strategic schedule to generate uncertainty based short-term plans. The objectives considered minimize deviations from the quality and quantity of ore to be delivered to the process plants, maximize equipment utilization, and maintain compliance with the strategic mine plan. Aligning the short-term plan to meet long-term strategic mine plan is critical. The operational simulation aspect of the proposed method accounts for shovel movements, haulage capacity, haulage profile, available trucks and dispatching efficiency, equipment failures, equipment availability, real time grade blending, and truck and shovel rate impacts due to their location. This model leverages historical production data to better predict haulage capacity based on the mining configuration and prevent overestimation of mine production. Similarly, the time-based fluctuations of material grades and plant feed can be considered on a finer timescale to ensure the plant is satisfied. The simulation optimization approach is used to generate a short-term production schedule based on single year of a long-term strategic mine plan considering five shovels, two crushers and a waste dump. The simulation was run for six months considering operational uncertainty. The optimizer chose to dispatch trucks to waste to ensure ore production was met and the model allocated more trucks to closer shovels to maximize production. A major limitation of this aspect is the short-term schedule ends up taking an efficient means of operating when optimizing the production schedule and fails to account for the repercussions of short-hauling material in early periods. This can lead to large increases in cycle times in later periods, perhaps requiring the purchase of additional hauling equipment to maintain expected production rates. This further underlines the need to integrate short-term production scheduling with the long term mine plan to maintain project value. The detailed approach highlights some important parts of the haul cycle, shovel assignment and blending requirements, but could also consider opportunities to integrate waste dump construction, rehabilitation and other environmental requirements. There is also opportunity to simulate crushing and processing behaviour to extend the model to the remainder of the mining complex and to integrate geological uncertainty.

Blom et al. (2017) suggest a framework that outputs multiple short-term production schedules for an open pit mine that extract ore and waste using different sequences and meet desired production targets. The method produces several short-term schedules that are then evaluated for quality and mineability. A single open pit mine is considered that generates one product with lower and upper bounds on the quality. The optimization has a constraint that forces the schedule to fully utilize available processing capacities and consider mining precedencies in the form accessibility. Multiple solutions are created, each providing a schedule and destination policy for material during the planning timeframe. A concurrent rolling time horizon-based algorithm aggregates time periods into groups greater than two, then solves each of the groups independently. Lastly, a split and branch approach allows the method to generate several schedules. The objectives are prioritized by their importance and the mixed integer programming approach sequentially tries to optimize each objective for each of the groups. Each objective is solved respecting the previous objective in the sequence by adding constraints of all those less than the previous solution. The mixed integer program is then solved considering the next objective. The advantage of this approach is the fast solve time. This comes at the cost of optimizing time frames independently, not considering the implication of past and future time periods and prioritizing objectives based on priority, and not considering the contribution to the objective function. Lastly, it forces the processing plant to run at full utilization which, likely reduces the number of available decisions and could lead to sub-optimal decision that impact long term value. The model also does not account for updating information and technical risk making the schedules both, sub-optimal and unlikely to be obtained.

L'Heureux et al. (2013) present a mixed integer programming model for short-term planning in open-pit mines. The scheduling approach considers the precedence of operational activities including drilling, blasting, and transporting material while accounting for typical accessibility constraints. In addition, the location of drills and shovels and access point are incorporated. A period of up-to 3 months is considered, which assesses the sequence of blocks to be mined along with equipment allocation. The approach limits the blocks to be mined based on the strategic mine plan and considers a much smaller number of blocks to plan the short-term schedule. The complexity of the short-term problem still leads to a very large model due to the additional operational details present. Each block is indexed by its block number and period to be extracted and for a schedule discretized by days over a three-month period it is clear that the number of variables is large. In addition, ensuring drilling patterns are grouped together for drilling and blasting, capacity and blending, and precedence constraints are considered the problem becomes very difficult to solve. The problem is reduced by removing several blocks in the geological model due to accessibility constraints. The rest of the blocks are grouped into areas by combined adjacent blocks to be mined during the scheduling process. This way shovel moves can be penalized for long-moves between areas whereas small moves within areas are penalized less. Groups of blocks are then considered to be blasted by clustering them into groups. Clusters can contain blocks in common. The model aims to minimise the costs of mining including moving shovels, drilling, and blasting while also including lower bound constraints on shovel utilization. This is the lower bound that is found when optimizing the long-term mine plan. The flexibility of this model allows for violations while, still trying to minimize costs and provide an effective short-term production schedule. The method considers a wide number of components in the mining complex; however, it does not consider aspects at the processing facility and further downstream. The optimization model is unable to reach an optimal solution within 10 hours. Therefore, a number of branching techniques including fixing certain variables to zero, prioritizing branching steps, eliminating symmetry between solutions, pre-assigning the shovels to locations, and adding a lower bound on shovel displacement are considered. The model was only able to be solved for 12 periods and a heuristic approach is suggested for future work. The mathematical framework adds several important operational aspects but fails to include geological uncertainty.

Mousavi et al. (2016a) present a short-term production scheduling approach that test three different metaheuristics: simulating annealing, tabu-search and a hybrid tabu-search simulated annealing. This method provides a solution approach for short-term production scheduling in large-scale real-world applications that have too many decision variables and constraints to be solved with commercially available solvers. The framework considers stockpiling, ore blending, mining width, and dynamic destination assignment. The two neighbourhood moves are time-based moves and destination-based moves. Time moves change the extraction period of blocks and are prioritized based on the periods with the minimum material being sent from mine to processor or the maximum flow from a mine to stockpile. A weighed sampling method is used to randomly select the period. A period either accepts new blocks from another period or transfers blocks to another period based on the capacity constraints. Destination moves are considered as a sub-problem with the objective of maximizing the movement of material from the mine to the process plant by minimizing rehandling from stockpiles. This is solved using mixed integer programming. Several different cooling schedules were tested for the simulated annealing approach and a tabu-search method uses a memory structure to store previously visited solutions making them tabu. This prevents revisiting the same solution again for a certain number of iterations. The solution quality was improved by using the hybrid tabu-search simulated annealing framework, particularly for larger instances highlighting the importance of memory-based methods.

Liu and Kozan (2019) present an initial development of combining strategic, tactical and operational scheduling approaches in a consecutive framework using a commercial solver (CPLEX). The integration of the three levels of planning allows for determining the appropriate size mining jobs and planning for the allocation of equipment at each stage is critical. They suggest adjusting the mine production schedule and equipment allocation repeatedly to achieve the appropriate balance between the number of available resources and the expected production. The drilling, blasting, and excavating activities are sequenced consecutively and when the current planning information can not be completed based on the available resources the production targets can be adapted. This method shows that interacting between different planning horizons allows for possibilities to reduce resources and increase efficiencies in a mining complex by minimizing the completion time and satisfying the precedence relationships between each mining task.

Matamoros and Dimitrakopoulos (2016) develop a short-term production scheduling approach using stochastic integer programming. Two major advantages of this approach over previously mentioned approaches are the mathematical formulation integrates fleet and mining considerations alongside the extraction sequence and actively manages supply and equipment uncertainty. Typically, the extraction sequence is optimized prior to the equipment assignment leading to sub-optimal solutions and each method is based off deterministic inputs failing to manage both the geological and operational risk of meeting production targets. The formulation aims to minimize the expected total mining cost along with deviations from production targets and mineable schedules. The first stage decisions determine if a block is mined in a sector, if a shovel is allocated to a sector, and the number of trips truck trips to each sector and shovel during each period. Next, the second stage recourse variables measure deviations from ore production, quality targets, hauling cost, mining width and equipment utilization. The mathematical model ensures feasible production schedules that account for shovel movements, haulage uncertainty, and blending requirements. The model does not explicitly try to maximize value by considering the financial impact of meeting these requirements and only minimizes mining costs. Additionally, this objective ensures that shovels maintain a certain level of productivity that may not be necessary to meet production targets leading to a conflicting objective between cost minimization and max utilization. The targets should be designed in a way to ensure the complete extraction of the scheduled areas in the long-term plan, while aiming to maximize value.

The short-term stochastic integer programming formulation is applied at a multi-element iron deposit with various quality constraints on phosphorous, silica, alumina, and loss of ignition. Additional short-term information is included in the formulation to account for hauling distance, hauling speed, and cycle times collected from operational data. The solution leads to a $15 M cost reduction when compared with the schedule that ignores geological and operational uncertainty. There are however several drawbacks to this approach. First, operational uncertainty is sampled from Gaussian distributions derived from historical data, which is not representative does not consider the joint uncertainty between components. Second, the shovel production is bounded by the worst-case scenario disregarding the value of the simulated equipment performance. Third, the mineability and accessibility constraints are highly inefficient. The formulation simultaneously optimizes the mine production schedule with mining fleet considerations. However, this development does not include additional operational aspects at the processing facilities apart from blending and requires a complete re-optimization based on any change to the operating environment. The pre-processing step only allows blocks that have the possibility of being used for blending that are between the allowable grades, not accounting for blending materials outside these boundaries potentially leading to lower valued schedules.

Quigley and Dimitrakopoulos (2020) extend the previous model described by considering multiple pits, processing streams and material types in a mining complex. The stochastic programming framework allows for location dependent shovel moves and a new approach for simulating equipment performance. In addition, horizontal block access constraints for facilitating smoother mining are developed. The precedence constraints promote mining spatially connected blocks in the same period to improve accessibility and reduce shovel moves. A pre-processing technique accounts for horizontal precedence based on the accessible direction that comes from a pre-designed haulage ramp. The stochastic programming formulation is solved using the branch and cut algorithm in CPLEX, which leads to inefficient run times. Therefore, a sliding time window and early start heuristic was introduced to solve the problem improving plan compliance and accounting for uncertainty.

Both and Dimitrakopoulos (2020) present a stochastic optimization approach that jointly integrates equipment and geological uncertainty. The stochastic mathematical programming model aims to simultaneously optimize the short-term production schedule with fleet management. This is compared with a conventional approach that optimizes each step independently by first optimizing the short-term production schedule and then allocating the mining fleet. This simultaneous consideration minimizes the number shovel moves and substantially reduces haulage requirements, leading to an improved short-term production schedule. Fleet management optimization typically aims to optimize the shovel locations in an open pit mine and determine the appropriate number of trucks for each shovel. Then, each truck is assigned to a destination. The method is formulated by maximizing the value of the products sold and does not consider the economic value of the block overcoming several limitations associated with the previous developments. This formulation is the first short-term formulation that directly accounts for the profits generated in the short-term scheduling problem and only aims to minimize deviations from shovel and truck production to ensure plan compliance. This allows the destination policy to be determined simultaneously during the optimization process accounting for blending from multiple sources, along with the newly incorporated decisions. The complicated problem is solved by leveraging an adaptive multi-neighbourhood simulated annealing approach that considers perturbation for extraction decisions, equipment assignment, and downstream material allocation.

Both and Dimitrakopoulos (2021; 2023a) integrate a throughput prediction model for the ball mill throughput into a simultaneous short-term production scheduling framework for mining complexes. Measurement while drilling data obtained off operational drills is used in combination with geological data to predict throughput of the mill overcoming the need to complete expensive test work. A multiple regression model is developed that considers the incoming feed production data to forecast ball mill throughput. The throughput rates are retrieved from the regression model as the schedule is modified in the simultaneous stochastic optimization framework to improve throughput accuracy and create more accurate and reliable production forecasts. A case study is completed at the Tropicana Gold Mining Complex where a multiple regression model is fit using a material tracking framework that determines the relationship between blended rock properties to observed ball mill throughputs. The results show that 7% of planned tonnage would end up becoming unprocessed in certain periods due to ball mill throughput bottlenecks whereas when ball mill predictions are integrated into production scheduling optimization framework there is only an expected shortfall of 0.3%.

Both and Dimitrakopoulos (2023b) integrate a geometallurgical prediction model for reagents and other consumables in a simultaneous stochastic optimization framework for short-term production scheduling. The prediction model considers the blended material attributes entering the processing facilities as the input variables and maps it to the expected consumption rates of reagents and consumables in an operating gold processing plant. The multiple linear regression prediction model is integrated into the simultaneous stochastic optimization framework for optimizing short-term production schedules to assess the impact of reagents and consumable consumption on short-term production schedules while also considering four open-pit mines and a stockpile facility. The integration of reagent and consumable prediction provides improved accuracy in expected consumption rates of processing reagents and consumables and leads to a 3.22% increase in profit.

1.2.4 Adaptive frameworks for optimizing mining complexes

Current state-of-the art methods used in machine learning and reinforcement learning applications have qualities that could be highly advantageous for solving short-term scheduling problems. First, these methods are data-driven and can work directly with the underlying data that impacts the decisions without relying on a specific model. Second, they are easily updatable in the presence of new-incoming information and have can adapt decisions in near real-time. Lastly, large stochastic programming formulations can likely be decomposed into smaller and potentially less complicated learning tasks (Bengio et al. 2021). For example, in a mining complex the learning tasks can be broken into determining the extraction sequence, fleet assignment, destination policy, stockpiling, blending, and operational alternative that maximize some global objective. Reinforcement learning approaches in particular provide a method to compute a policy that aims to maximize cumulative reward and is improved through continuous trial-and-error. Adaptive state-dependent policies are desirable in production planning because as new information becomes available there is an opportunity to update the short-term production schedule to better meet the long-term production schedule. Additionally, the optimized policies learned in previous interaction between the long and short-term production schedule can potentially be made general enough to apply in the presence of additional information. As planning processes currently work, the short-term scheduling of mining complexes occurs much more frequently than long-term planning. Therefore, long-term schedules should be adapted to changes in the short-term schedule and vice-versa. In this section, a few achievements in adaptive policies and reinforcement learning frameworks in mining will be discussed along with a number of algorithms in reinforcement learning that could be utilized to connect short and long-term planning horizons and link production schedules of different timescales.

Paduraru and Dimitrakopoulos (2018) present an approximate dynamic programming approach that generates an adaptive policy for the short-term optimization of material flow in a mining complex. The method uses a state-dependent policy that takes new incoming information as it becomes available to make adaptive decisions for the materials to be mined and processed. These robust policies can be applied across different timescales leading to an opportunity to pass information from one-time scale to the next. This a critical step to leveraging new information and ensuring the production schedule is optimized based on the most up-to-date information. The framework optimizes the flow of materials as new information becomes available about the extracted material. The state dependent policies are a function of several different properties within the system and can be obtained at a more frequent resolution by observing the state of the mining complex at any given point of time. The state of a mining complex in this context relates to a numerical representation of the key components of the mining complex, such as, the material grades, tonnage, and number of secondary elements. In addition to the material properties, the status of various components such as stockpile inventory, mill feed grade, and product available for shipment are some examples of information that can be taken from the mine production schedule. The state of the environment is updated when new information becomes available and this can be in the form of grade control, sensory data from stationary and mobile equipment, market price forecasts and more. To optimize the state dependent policy, in the example here a material destination policy, a parametrized policy is created that increases efficiency and minimizes mill disruptions. The optimization of the three parameters required for determining the destination is completed using Bayesian optimization a technique that maximizes the average posteriori improvement in the objective function for determining the optimal cut-off grade. An application of the proposed approach is completed at a mining complex with a single mine and six destinations. The objective function aims to determine a state-based destination policy for each block that maximizes the net present value of the operation. By learning a state-dependent policy, this work shows that the policy is capable of responding to new information in a very quick and efficient manner, which is purely data driven. The challenge remains in how practical is a black-box approach that would require a computer program to calculate the best policy based on the present state of the operation to predict the best destination. Lastly, the framework solely focuses on determining the short-term destination policy in a mining complex and requires a predefined extraction sequence leading to a locally optimal solution.

Paduraru and Dimitrakopoulos (2019) propose a policy gradient reinforcement learning algorithm to optimize the destination policy in a mining complex by optimizing the parameters of a neural network and accounting for supply and equipment uncertainty, instead of using an approximate dynamic programming approach. This manuscript showcases the advantage of reinforcement learning over other optimization approaches based on their ability to generate policies that react to a wide number of situations making them adaptive in nature. This is different from standard stochastic programming techniques where each time the state of the mining complex changes due to new-incoming information a new optimization must take place. The policy gradient approach discussed considers more attributes in the mining complex and uses discrete event simulation to model the material flow up to the crushers under equipment uncertainty. The destination policy is a mathematical function mapping the state of a mining complex to the appropriate decisions using a neural network architecture. The resulting policy that is computed does not require re-optimization and can be reused for future decisions even if new information unfolds. A practical aspect of this type of framework is that the policy can be applied to rapidly determine the best destination given the current state of the mining complex and could potentially be integrated into the simultaneous stochastic optimization of mining complexes similar to the prediction models discussed in Both and Dimitrakopoulos (2023b; 2023a). The reinforcement learning method accounts for modelling the joint uncertainty distribution between equipment and orebody scenarios, however, this is accomplished using a generative model meaning that a single geostatistical model is sampled and then the equipment uncertainty model is generated. The challenge here remains a geostatistical problem associated with using a single stochastic orebody realization and optimizing using this representation of the deposit. The local uncertainty and variability of a single realization of the mineral deposit is not representative of the spatial uncertainty between blocks and the joint variability within the volume to be mined must be considered to properly describe the characteristics of the material supply. The policy gradient approach increases expected net present value by 6.5% when compared to destination policy in place at a copper mining complex. The major limitations of this approach are the method only accounts for a single element copper and the production schedule is not predetermined.

Kumar et al. (2020) present an adaptive framework for updating short-term scheduling decisions in a mining complex considering multiple elements and extending the policy gradient method described previously. A continuous updating framework is developed that uses policy gradient reinforcement learning to optimize the production schedule along with an extended ensemble Kalman filter for updating the equipment and geological uncertainty models based on new-incoming information. New information in the form of blasthole sampling is used to update the uncertainty models and the schedule is then adapted to this new information. A neural network reinforcement learning approach is then trained using machine learning techniques by following the policy gradient and maximizing the future return. The learned neural network policy is used to adapt the short-term production schedule. The updating scheme uses an extension of the ensemble of Kalman filter that can update multi-element mining complex. In addition, a hauling and extraction simulator is used to generate samples for training a deep-neural network. A large hurdle is overcome by combining an updating framework with the optimization of the destination policy in a single framework. The adaptive framework is tested in a copper mining complex and adapts the flow of materials using new-incoming information. The proposed framework is able to increase the total cash flow in an operating mining complex by adapting the destination policy to new information. The major limitation of this approach is once again only the destination policy is optimized and a predefined extraction sequence is used leading to a sub-optimal solution.

Kumar and Dimitrakopoulos (2021) expand the previous framework to simultaneously determine the short-term extraction sequence and a destination policy in a mining complex. A self-learning reinforcement learning agent is implemented using a Monte Carlo search trees to determine an optimal policy for material extraction and destination that adapts to new incoming information. The framework is based on AlphaGo and AlphaGoZero algorithms (Silver et al. 2016; Silver et al. 2017). A policy is learned to determine the probability of taking an action and the value function is also learned to estimate the value of the short-term production schedule from a given state. The short-term production schedule is optimized in a sequential process that takes the current state of a mining complex and selects an action based on the probability of each action output from the neural network policy. After each action is taken a reward is received, the new state of the mining complex is unveiled, and the next action is taken until the short-term production schedule is determined. New information is collected from sensors and used to update the supply and equipment uncertainty with Kalman filter and Monte Carlo simulation, respectively. The updated uncertainty models are included in the new state of the mining complex. The adapted production schedule is able to meet long-term production requirements and further utilize processing capabilities leading to a 12% increase in profit over a 13-week planning horizon. The major limitations of this method are the framework fails to consider stockpiles and only accounts for the mill destinations simplifying the problem substantially. In addition, the policy is based on a very small action space that only considers adding block-by-block evaluations. This can become extremely cumbersome when training the model as each time a Monte Carlo tree search is performed every block must be visited. In addition, the shovels were preassigned to a set of 81 blocks meaning shovel movements and access are not considered. The method effectively determines the extraction sequence and destination policy for a 13-week time horizon in a reasonable amount of time (2 min) after 2 days of training.

A few state-of-the-art methods in reinforcement learning for scheduling and solving combinatorial optimization problems will be discussed briefly based on the opportunity to further develop the adaptive frameworks discussed and find strategies to connect short and long-term planning horizons. Multi-agent reinforcement learning allows agents to use common knowledge to execute complex coordinated policies with actor-critic reinforcement learning (de Witt et al. 2020). The agent is the learner or decision maker that interacts with its environment to achieve a goal based on some cooperative reward. Here, a master problem is created where higher level agents coordinate groups of agents by conditioning on their common knowledge. Therefore, agents can share observations, parameters, and gradients, however, the agents can still make their decisions based on their independent observation of the state. The idea between common knowledge agents is that complex coordination policies can be creating by exploiting knowledge between groups of agents using a form of hierarchy that can be described using a policy tree. Another example of multi-agent frameworks that leveraged to solve a combinatorial optimization problem using metaheuristics includes the work by Silva et al. (2019). Each agent takes actions independently and shares information with each other through the environment. The agents then learn from their actions based on the experience interacting with both other agents in the environment. The interaction between agents provides a global understanding of the problem that can influence the quality of the solution as seen in the simultaneous frameworks discussed previously. However, the challenges with these methods are they remain unstable at scale, and this relates to unstable agents that begin with a randomly initialized value and policy networks. In addition, in mine production scheduling each agent has a very different workload as the extraction sequence and destination policy require a larger search for optimal policies. A major area of research is to create agents that can quickly learn the optimal policies for production scheduling decisions mining complex that obtain stable results.

1.2.5 Modelling supply uncertainty

Stochastic optimization frameworks require a set of equiprobable stochastic orebody simulations of the material supply. These realizations can be generating using a wide range of geostatistical methods. An exhaustive literature review is extensive, therefore, to highlight key developments we will briefly describe the limitation of traditional two-point geostatistical simulation methods and the advantages of high-order statistical methods. Lastly, an investigation on updating frameworks for geostatistical simulations.

Geostatistical simulation methods are based on the concept of random fields and random functions. In mining applications, it is critical to understand the joint uncertainty of the attributes of interest at various locations in a mineral deposit. For example, the probability of having connected high-grade values in space is indicative of whether the area should be mined based on the uncertainty and local variability in the material grades. This type of input is valuable as stochastic optimization frameworks can manage the risk of supply uncertainty directly in their formulation by accounting for variations in the connectivity of high grades and their locations in the mining complex by observing the spatial-temporal relationship conditional to the extraction sequence.

Traditional estimated orebody models are generated by a number of methods, such as kriging (Isaaks and Srivastava 1989; Goovaerts 1997; Rossi and Deutsch 2014), to estimate the attribute of interest over the study area. An estimate for each location to be predicted is taken independently of neighbouring estimates with the ultimate objective to minimize the local error variance to form the best linear estimator. These methods tend to smooth local variability and misrepresent the proportion of materials in the ground. Typically, small values are overestimated and large values are underestimated (Goovaerts 1997). The smoothing effect reduces ones ability to detect patterns of extreme attribute values. The production schedule and material grades drive the profitability of a mining complex making it vital to reproduce the distribution of high-grade valuable material in space along with the connectivity and structure of the mineralization. Lastly, the data configuration impacts the effects of smoothing where smoothing is minimal close to the sample data and increases as it gets further away from these locations, potentially leading to artificial structures.

Stochastic orebody simulations can be used instead of an estimated model to represent the uncertain material supply. This enhances the understanding of the uncertainty and local variability of the minerals in the ground. A group of equally probable stochastic realizations map the uncertainty and local variability of the material supply and reproduce the statistics of the sampled exploration data. The data values are honoured at their locations and the mean and variance are reproduced; this includes the histogram of the sample data. Lastly, the covariance or variogram model reproduce the short-scale variability of the attributes under study. The most popular simulation approaches in industrial applications are sequential simulation approaches, which apply the recursive application of Bayes' rule to decompose an N-point conditional cumulative distribution function as product of N one-point conditional cumulative distribution functions.

Sequential Gaussian simulation (SGS) is one of the most popular simulation approaches and assumes a multiGaussian random function model (Isaaks 1991; Goovaerts 1997). The random function can be fully parametrized by the mean and variance. However, the major limitation of this approach is the multiGaussian random functions only account for first and second order spatial statistics making it unable to model complex spatial phenomena, such as, curvilinear structures and non-linear features. In addition, multiGaussian random fields are naturally high entropy, which does not allow for significant spatial correlation between extreme values and forbids geological organization (Journel 2002). An alternative method is sequential indicator simulation, a non-parametric and non-Gaussian approach (Journal and Alabert 1989; Goovaerts 1997). The indicator approach provides the possibility to account for spatial continuity by generating a number of indicator variogram models and provides an approach for introducing soft information. The indicator approach requires multiple variogram inference; a tedious task and suffers from order relation problems that can happen due to negative kriging weights since the conditional cumulative distribution functions are not solved jointly. In addition, once again this method only considers up-to two-point statistics, only guaranteeing the reproduction of the mean and variance. A notable example, shows that realizations generated from three different simulation approaches can have drastically different geological representations while reproducing the same variogram and two-point statistics proving a major limitations of two point approaches that ignore higher order statistics (Journel 2002). The variogram or covariance model is not enough to characterize even the simplest of curvilinear structure and alternative methods are required that utilize higher order statistics.

High order statistics – moment and cumulants – account for different geometric data configurations by using two or more separation vectors. This helps reproduce the connectivity of high grades and low grades in a mineral deposit. de Carvalho and Dimitrakopoulos (2019) present an example that showcases the impact of simulating realizations using a high-order simulation approach at a mining complex. The simultaneous stochastic optimization framework compares optimizing a mining complex using a set of simulations which, reproduce the connectivity of high and low grades within the mineral deposit and compare it to a two-point simulation approach that only guarantees reproduction of up to two-point statistics. The production schedule generated using a set of orebody realizations that characterize the non-linear spatial connectivity drastically changes the life of mine production schedule due to the consideration of high order statistics in the simulation framework. In addition, the high order simulation approach is able to improve the ability to capture the spatial connectivity of the mineral deposit by accounting for information of higher orders. The ability to effectively model the geological phenomena understudy is critical as all future mine planning decisions are generated based on this assumption.

Dimitrakopoulos et al. (2010) introduce a set of definitions of non-Gaussian spatial random functions and their high-order spatial statistics. In particular, they introduce the use of spatial cumulants as a possible direction for a high-order geostatistical simulation approach. Spatial cumulants of higher order are shown to capture the multiple-point periodicity, connectivity of extreme value and the spatial representation of the geological phenomena under study. Cumulants are proposed as a method to capture information associated with the spatial properties of a mineral deposit that accounts for multiple points. A unique feature of cumulants is that they are function of the moments of lower or equal orders; therefore, cumulants provide a robust mathematical framework for reproducing statistics across low and high orders. Lastly, unlike multiple point simulation approaches, not described here, they are based on a consistent mathematical formulation and are not strictly driven by a training image (geological analogue) and their associated algorithms, see (Strebelle 2002; Zhang et al. 2006; Wu et al. 2008; Mariethoz and Caers 2014).

Mustapha and Dimitrakopoulos (2010) present a sequential simulation algorithm that accounts for information that is inferred from a training image and production data using Legendre polynomial and spatial cumulants. The method is able to accurately reproduce the statistics up-to order five on a sparse data set. The algorithm pre-processes the spatial cumulants found in the training image and data. Afterward a sequential simulation framework is used to simulate accounting for these relationships. The coefficients of the Legendre polynomials are calculated from the cumulants, accurately capturing complex geological patterns. The limitations of using Legendre polynomials as an orthogonal basis include sensitivities to rounding errors and they may not always converge to an analytical function. In addition, the preprocessing step requires a predefined spatial template. (Minniakhmetov and Dimitrakopoulos 2017b; 2017a) and (Minniakhmetov et al. 2018) extend this framework to jointly simulate multivariate deposits and categorical data and improve on the basic limitations using a dynamic spatial template for finding replicates. The method uses a high-order simulation approach that approximates high-order spatial indicator moments. A piecewise polynomial function connects pieces under smoothness conditions. The locations where these pieces are meet are called knots and knots impact the flexibility of approximating particularly when there are discontinuities. To maintain an orthogonal basis, Legendre like splines and a linear combination of B-splines are used to improve the approximation of the conditional cumulative distribution function. The main limitation of this approach is it can overfit if using a high number of knots and requires more computational effort for calculating the Legendre coefficients. In the multivariate version, for simulating several material attributes, diagonal domination is used as a decorrelation step to obtain independent non-Gaussian factors. These factors are then simulated independently and reconstructed. The non-diagonal terms of the covariance-variance matrix are minimized, and the diagonal values are maximized to minimize dependence between factors. (Yao et al. 2018) improve computational efficiency by using spatial Legendre moments and an empirical formulation eliminating the need to explicitly compute high-order moments or cumulants using a recursive formulation.

High-order simulation approaches are able to reproduce non-Gaussian and non-linear spatial patterns an important property related to the geological properties in most mineral deposits. More realistic realizations can be obtained by accounting for the high-order statistics and overcoming the limitations of high entropy approaches. The benefit of adequately accounting for the connectivity of high grades using high order statistics leads directly to different optimal production schedules.

A challenge with all simulation approaches is the simulation process is a time-consuming task. However, it is an extremely valuable investment for creating a risk robust production schedule and achievable financial forecasts that account for uncertainty. Recent, high-order methods are able to better capture the spatial uncertainty and local variability in the mineral deposit by accounting for the connectivity of high and low-grades. To overcome the limitation on the time required for simulation approaches updating frameworks are required for integrating frequently gathered new information. This also allows for adaptive decision making to be made under uncertainty using the most recent information available.

Vargas-Guzmán and Dimitrakopoulos (2003) propose a method for conditional simulation with successive residuals for stationary and ergodic Gaussian random functions through a column wise decomposition of the covariance matrix. The successive process can be thought of as a method to separate the influence of different data sources allowing recalculation of only sources that allow for new information when updating. Jewbali and Dimitrakopoulos (2011) leverage this property to allow for the rapid updating of simulated realizations. However, the major limitations of this method are the location of future data must be known and predefined and the covariance matrix must be stored in memory.

In a number of industries, data assimilation techniques are applied to update a prior model using new incoming information. It is a widely used technique for reservoir forecasting, atmospheric and oceanographic sciences and more recently resource model updating. Typically, there are two sources of information, both are imperfect and can have inaccuracies. This includes a model, a predicted numerical representation, and observations from sensory data that can have errors. The goal in data assimilation is to combine these two types of data to improve prediction. The use of data assimilation techniques to continuously update the model-based predictions with observations gathered during production can improve ones understanding of a mineral deposit. If acted upon, this can improve operational decisions. An early approach for data assimilation is Kalman filter, which is optimal under several assumptions (Kalman 1960). First, the prior distribution is Gaussian, the observation operator is linear, and it is Gaussian in the observation error. A regression coefficient is used that takes a linear combination between the prior model and the observations weight by the Kalman gain based on the covariance matrix and the measurement error. The observations likely are also uncertain. The optimal posterior estimate of uncertainty model is then obtained by calculating the difference between the predicted value and the observation and the Kalman gain governs the impact of the newly observed value on the model prediction. The Kalman gain is computed to minimize the a posteriori error covariance.

The Kalman filter has been extensively studied and can be extended to the extended Kalman filter, ensemble Kalman filter, and many more. The extended Kalman filter is an approximate filter for non-liner systems aims to linearize the model around the current estimate (Jazwinski 1970). The extended Kalman filter approach in this case is no longer an optimal estimator and suffers from divergence due to the linearization. On the other hand, ensemble Kalman filter explicitly replaces calculating the covariance matrix by an estimation of the covariance from a set of ensemble members (sample covariance), allowing the model to be used on larger problems where calculating the covariance for high-dimensional problems is infeasible (Evensen 1994). An extensive review of ensemble Kalman filter in reservoir models is well documented for petroleum models (Aanonsen et al. 2009). The Ensemble Kalman filter is limited to updating parameters with linear statistical relations and treats the observations as Gaussian random variables. Since higher moments are excluded, the model tends to remove non-Gaussian random functions. In addition, the size of the ensemble needs to be large enough to correctly approximate the covariance. Small ensembles can lead cross-covariance having undesirable correlations and can lead to inbreeding. Non-linear problems have been tackled using ensemble Kalman filter in a number of ways. Chen et al. (2009) emphasize the challenge of using Ensemble Kalman filter when the model parameters, state variables and observations are strongly non-linear through a change in parametrization with alternative variables. The Kalman filter approach has been developed and is well adopted for dynamically updating reservoir models and mining models, however, it suffers from the Gaussian assumption and considers up to two-point statistics, the first two moments. The Gaussian assumption makes the updating framework simple to integrate however, when applying on non-Gaussian variables it can lead to strong violations. The non-Gaussian characteristic of the Kalman filter can come from both the probability and the model dynamics.

Benndorf (2014) introduce a sequential updating framework for mineral resource modelling using available sensory data. The challenge with the updating framework is material is often sensed based on a blend of material coming from separate locations. The method adapts the Kalman filter and ensemble Kalman approach to account for this challenge by tracking the blended material back to the source locations for updating. Benndorf and Buxton (2016) suggest integrating the updating framework described to update resource models and then rapidly optimize the short-term production plan. They also discuss the impact of multiple sources of material being mined and collecting data for updating. Wambeke and Benndorf (2017) assimilate online production data from sensor-based information into the grade control model. A computationally efficient method is developed by excluding the forward observation model and building a forward simulator independently from the existing code that translates grade control realizations into observation realizations. This way empirical covariances can be calculated and a linear forward simulator is not required. In addition, the method manages non-Gaussian distributions by applying a normal score transform. Wambeke et al. (2018) demonstrate the real-time update of a geometallurgical model by improving the predictions of the spatial bond ball mill work index estimates by calculating deviations between the predicted and actual mill performance and transferring these details to the geometallurgical model. Lastly, Kumar and Dimitrakopoulos (2021) extend the ensemble Kalman filter framework to update multiple correlated variables.

These approaches are limited by the Gaussian assumption and can be used for continuous model parameters. There is no guarantee that the data assimilation approaches preserve the geological consistency of high order statistical measures. Although, several methods have attempted to account for this by applying various transforms (Kumar and Srinivasan 2019; 2020). As discussed, accounting for high-order statistics is an effective approach to simulate highly complex non-linear and non-Gaussian geological phenomena to maintain the connectivity of high and low-grades and how they behave within a mineral deposit. In addition, categorical measurements of the contact of different material types with properties that impact waste management such as, overburden contact and acid generation. These updating frameworks are limited in their ability to account for high-order spatial statistics and future research is required.

1.3 Research goal and objectives

The goal of this thesis is to develop a method for connecting short- and long-term planning horizons in mining complexes with a simultaneous stochastic optimization framework that uses reinforcement learning to align and optimize production schedules at different timescales. This is to be accomplished while accounting for supply uncertainty. In addition to connecting planning horizons, several additional aspects of a mining complex will be extended to account for preconcentration techniques, waste management and reclamation to improve environmental performance.

Outlined below are the proposed objectives to reach this target:

(1) Complete a literature review of critical developments in simultaneous stochastic optimization and adaptive learning methods for optimizing mining complexes.

(2) Develop an infill drilling reinforcement learning optimization framework that links the value of additional information in mining complexes to long-term production scheduling decisions.

(3) Develop a reinforcement learning approach for optimizing short-term production scheduling decisions in mining complexes that leverages learned policies to solve a stochastic programming formulation that integrates preconcentration.

(4) Embed the short-term reinforcement learning framework in (3) with a long term simultaneous stochastic optimization framework for mining complexes to develop a formulation that connects short- and long-term production planning timescales.

(5) Extend the simultaneous stochastic optimization formulation for mining complexes to incorporate critical environmental considerations that actively improve waste management and reclamation practices in mining complexes.

(6) Summarize developments and provide guidance for future research.

1.4 Thesis outline

This manuscript-based thesis is organized into the following chapters:

Chapter 1 provides a literature review of the key topics that relate to the goals and objectives of this thesis. This includes developments leading to the simultaneous stochastic optimization of short- and long-term production schedules, models for incorporating preconcentration techniques, critical considerations for reclamation and rehabilitation of mining waste dump facilities, impacts of acid mine drainage, and adaptive learning techniques for optimizing policies with reinforcement learning.

Chapter 2 extends the simultaneous stochastic optimization of mining complexes to quantify the value of collecting additional information via infill drilling. A reinforcement learning framework is developed for selecting infill drilling locations that assesses the impact of additional drilling with stochastic simulations. The optimization approach learns to plan infill drilling locations that improve long-term production scheduling decisions by increasing value and mitigating risk.

Chapter 3 develops a novel simultaneous stochastic optimization framework for short-term production scheduling that uses stochastic programming techniques and actor-critic reinforcement learning with a discrete-continuous action space. The reinforcement learning agent learns to determine short-term extraction sequences given available access points and the uncertainty and local variability of materials. Practical schedules are obtained with a discrete-continuous action space used for extraction sequence decisions. The stochastic programming formulation is capable of optimizing mining complexes with multiple mines, stockpiles, and processing facilities. In addition, the method integrates preconcentration facilities into the optimization formulation allowing for the separation of coarse and fine materials, which improves feeds grades and reduces energy and water usage by rejecting coarse waste material prior to processing. The method is tested in a large-scale copper mining complex.

Chapter 4 develops a new optimization framework that connects planning horizons in mining complexes by simultaneously optimizing short- and long-term production schedules. The optimization formulation is solved with a combination of metaheuristics and an embedded reinforcement learning agent developed in Chapter 3. A metaheuristic solution approach modifies the long-term production scheduling decisions. Then, given the current long-term schedule the embedded reinforcement learning agent rapidly assesses the short-term production schedule for the first production year and communicates its findings. The method is tested in an operating copper mining complex.

Chapter 5 develops a stochastic programming formulation for simultaneously optimizing long-term production scheduling decisions with integrated waste management and reclamation considerations. Acid rock drainage is minimized by blending waste materials with different geochemical properties to create a waste product that acts as if non-acid generating by modifying scheduling decisions including the placement of waste rock material. In addition, progressive reclamation decisions are made to directly place topsoil on completed waste dump areas allowing for revegetation and improving environmental performance. A case study is completed at a multi mine copper-gold mining complex.

Chapter 6 presents the conclusions of this thesis and suggests possible future research opportunities.

1.5 Original contributions

I. **Development of a reinforcement learning based framework for infill drillhole selection in mining complexes that determines the value of additional drilling information by simultaneously optimizing a mining complex in light of new information that is sampled from a set of stochastic simulations of future infill drilling data.**

Past optimization frameworks for optimizing infill drillhole selection in mining complexes focus on minimizing estimation variance, which is influenced largely by geometry of the drillhole configuration. Additionally, infill drilling approaches that use stochastic simulations to assess future infill drilling information use indicators that compare block material classifications, which are expected to influence production scheduling decisions. However, selecting drillholes based on these criteria does not guarantee that it will impact the resulting production schedule and this information could instead be collected during grade control. The proposed method provides a novel framework that directly assesses the impact of future infill drilling information on long-term production scheduling decisions and seeks infill drilling locations that will provide information that can improve operational guidance by adapting the long-term production schedule to new information.

II. **Development of a simultaneous stochastic optimization framework with integrated preconcentration for short-term production scheduling that utilizes reinforcement learning with a hybrid-discrete continuous action-space and stochastic programming to solve a newly proposed formulation.**

Combining reinforcement learning with stochastic programing approaches for optimizing short-term production schedules has received limited attention in previous works. Adaptive reinforcement learning frameworks have demonstrated useful approaches for optimizing a subset of mining complex decisions, however, the applications typically do not account for critical operational aspects including stockpiles, preconcentration facilities and mining accessibility requirements. The proposed framework leverages reinforcement learning to determine the direction and quantity of material mined based on the distribution of material grades surrounding each mining area. An operational short-term production schedules is obtained by accounting for accessibility and optimizing critical downstream components including process plants, stockpiles, preconcentration facilities, and waste dump facilities. Grade-by-size material characteristics that influence the performance of preconcentration facilities are simulated to describe the behaviour of material at screening facilities, which can decrease the amount of waste entering the processing streams.

III. **Connected planning horizons by jointly optimizing short- and long-term production schedules to ensure compliance between schedules of different timescales and improve the accuracy of production forecasts.**

Existing methods for optimizing short- and long-term production schedules are typically optimized from the long-term to the short-term with an increasing level of detail being considering. The limitation of this approach is that the optimization of short-term production schedules is bound by the decisions made for long-term production timescales. The proposed approach aims to jointly optimize short-term production schedules of different timescales to ensure compliance between the production schedules and create more realistic long-term production forecasts that account for critical short-term considerations.

IV. **Integrating waste management and progressive reclamation into the simultaneous stochastic optimization framework for mining complexes.**

Waste management and reclamation are critical environmental requirements in operating mining complexes and require careful planning to minimize damages and return the environment to a productive post mining state. Past works fail to consider the uncertainty and local variability of geochemical properties that can cause acid rock drainage in waste dump facilities. In addition, reclamation is usually planned after long-term production scheduling decisions are made and uncertainty related to the construction material used for reclaiming waste dump facilities is ignored. A novel stochastic programming formulation for optimizing mining complexes with waste management and progressive reclamation is developed to overcome these limitations which, mitigates acid rock drainage and improves environmental performance by promoting early reclamation of waste dump facilities. Metaheuristic techniques for optimizing large-scale mining complexes under uncertainty have been significantly improved over the years. However, the selection of perturbations that are applied to modify the production scheduling decisions are often predetermined or randomly selected. A novel approach is proposed that uses contextual information from the optimization formulation itself and the past modifications made to select which perturbations to apply in an environment with competing alternate choices. The improvement obtained by applying different perturbations that modify production scheduling decisions are only partially known and are constantly changing through the optimization process.

2 A reinforcement learning approach for selecting infill drilling locations considering long-term production planning in mining complexes with supply uncertainty

In strategic mine production planning under uncertainty, the long-term production schedule defines the extraction sequence, destination policy, and processing stream decisions that maximize net present value and minimize technical risk. Optimized scheduling decisions are driven by the inputs, a set of stochastic orebody simulations that describe the uncertainty and local variability of the materials in the ground. Collecting additional information related to the deposit can be advantageous for improving long-term production scheduling decisions, as this information can be used to update the input stochastic orebody simulations. However, determining where to drill and whether it is worthwhile to invest in additional drilling is an extremely challenging decision with uncertain outcomes. This chapter of the thesis provides a novel approach for planning infill drilling by considering its impact on long-term production scheduling decisions. Actor-critic reinforcement learning is applied to select infill drilling locations and then using stochastic simulations and a simultaneous stochastic optimization framework the value of the additional information is quantified. Different from previous frameworks, the method seeks to optimize infill drillhole planning by determining drilling areas that are likely to improve the long-term production and financial forecasts.

2.1 Introduction

A mining complex is an integrated supply chain designed to transform extracted materials from several mines into a set of valuable products for delivery to customers and the market (Pimentel et al. 2010; Dimitrakopoulos and Lamghari 2022). Production scheduling decisions are optimized to maximize net present value and, in cases where uncertainty is considered, minimize technical risk related to the uncertain material supply. Recent advancements in stochastic mathematical programming have demonstrated the benefits of simultaneously optimizing long-term production scheduling decisions within a mining complex, while considering supply uncertainty (Montiel and Dimitrakopoulos 2015; 2018; Goodfellow and Dimitrakopoulos 2016; 2017). The production scheduling decisions are optimized simultaneously given the model of a mining complex and the various behaviour of the components within. Furthermore, the uncertainty and local variability of the material supply is accounted for and managed in the optimization framework by using stochastic orebody simulations to represent the quantity and quality of materials in the ground (Goovaerts 1997; Remy et al. 2009).

Infill drilling is important for capturing additional information about the mineral deposit that can potentially reduce sources of error by identifying boundaries and discrete ore zones to more accurately classify the mineral resources and reserves (Parker 2012). Planning which areas to invest in infill drilling remains a major challenge in the mining industry because the outcomes of drilling are uncertain and the information collected may not necessarily cause a change in the long-term planning decisions. However, infill drilling in areas that provide valuable information for orebody modelling including geostatistical simulations can directly impact the outcome of the resulting long-term production schedule, potentially leading to improved production forecasts. To assess the influence of infill drilling on production scheduling decisions, a new framework is developed herein that quantifies improvements to the long-term production schedule based on infill drilling data. This provides a new criterion for infill drillhole selection. Future infill drilling data is sampled from a stochastic realization and used to update a set of stochastic orebody simulations (Jewbali and Dimitrakopoulos 2011; Peattie and Dimitrakopoulos 2013). The production schedule is then reoptimized considering new information to account for the forecasted improvements to the long-term production schedule.

Past work aiming to optimize infill drilling consider minimizing estimation variance (Diehl and David 1982; Gershon et al. 1988; Soltani and Safa 2015), which is influenced by the geometry of the drillhole configuration, sample density, number of samples, and covariance structure (Cressie 1991). This approach ignores the critical impacts related to uncertainty and local grade variability of the attributes of interest (Ravenscroft 1992; Delmelle and Goovaerts 2009). Furthermore, additional drilling is often assumed to reduce grade uncertainty, however, Goria et al. (2001) and Dimitrakopoulos and Jewbali (2013) document cases where collecting additional information can lead to greater grade variability and uncertainty. This supports the suggestion to consider the potential impacts of infill drilling on the long-term mine plan. Several approaches for optimizing infill drilling decisions are founded upon geostatistical simulation techniques and have been introduced to overcome these limitations by accounting for the effects of different drilling patterns.

Dirkx and Dimitrakopoulos (2018) apply a multi-armed bandit approach for selecting infill drilling patterns in a stockpile using a criterion that accounts for changes to material classification. Infill drilling patterns that cause the largest changes to the material classification of stockpiled material are selected as this is expected to reduce misclassification of these materials. To assess the value of infill drilling, additional data is retrieved from a stochastic realization of future infill drilling data and used to re-simulate a set of stochastic orebody simulations. Boucher et al. (2005) use geostatistical simulations of correlated variables to observe the impact of drilling on the forecasted profits obtained while evaluating different stockpiling options. By virtually drilling a simulated realization of the future infill drilling data (assumed to be the true deposit) and using the information to re-simulate the deposit with denser drilling patterns, an optimized configuration is determined by comparing block classifications and economic indicators. Kumral and Ozer (2013) propose an approach that aims to optimize the drilling configuration with a genetic algorithm that minimizes the maximum interpolation variance by sampling the drilling information using future simulated drillhole data and available exploration data. A limitation of these simulation-based methods is that they do not consider the relationship between the drilling information collected and potential effects in long-term mine planning and production scheduling, which could potentially lead to additional value.

The approach presented herein introduces a new framework for optimizing infill drilling decisions that considers long-term production scheduling decisions in the context of the simultaneous stochastic optimization of mining complexes by evaluating potential improvements to the long-term production schedule and associated forecasts. Actor-critic reinforcement learning with deep neural networks (Lillicrap et al. 2015; Sutton and Barto 2018) is applied to learn a policy for selecting infill drilling locations. The reinforcement learning policy guides the optimization process to relevant drilling locations with a stochastic search, learning the areas that lead to the largest improvements in the production schedule through continuous trial-and-error. Stochastic orebody simulations are updated with ensemble Kalman filter (EnKF), which overcomes the requirement to re-simulate the entire deposit each time new additional information is retrieved (Evensen 2003). Then, the long-term production schedule is simultaneously optimized considering the additional drilling information collected. Several recent works have demonstrated the successful performance of deep neural networks for a variety of optimization problems (Lillicrap et al. 2015; Mnih et al. 2015; Silver et al. 2017; Kumar et al. 2020; Kumar and Dimitrakopoulos 2021). Applying reinforcement learning for selecting infill drilling locations is advantageous as the stochastic search utilizes contextual information from the mining complex to guide the optimization process. Additionally, actor-critic methods work with continuous actions allowing the agent to select coordinates for drilling. The value of infill drilling in a mining complex is strongly related to the value of information, a frequent topic in decision theory (Yokota and Thompson 2004; Bratvold et al. 2009). Froyland et al. (2018) discuss how additional infill drilling relates to these topics and highlights that the value gained from infill drilling can only be realized if it results in changes to production scheduling decisions. However, a method was not employed to address these challenges.

In the following sections, the proposed method based on actor-critic reinforcement learning and simultaneous stochastic optimization is introduced for optimizing infill drilling that provides potential to improve production and financial forecasting in mining complexes. Next, a case study at a copper mining complex demonstrates the key aspects of the proposed approach and the impact of selecting a set of infill drilling locations. Conclusions and recommendations for future work follow.

2.2 Planning infill drilling with reinforcement learning and simultaneous stochastic optimization

The framework developed herein for planning infill drilling in mining complexes is outlined in this section. An overview of the workflow developed for infill drillhole selection is shown in Figure 2.1.

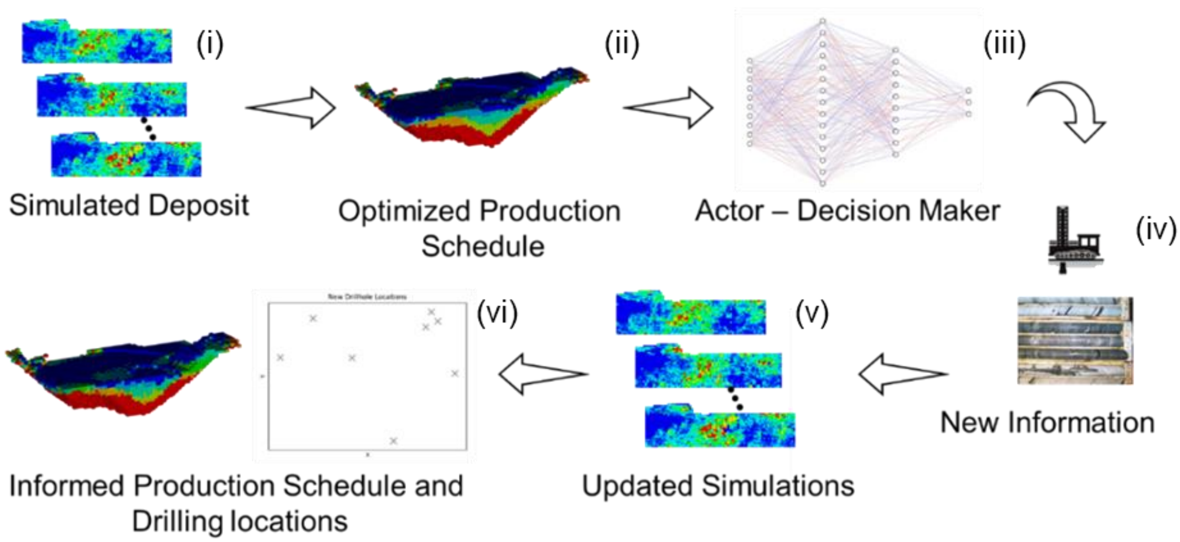


Figure 2.1. Primary workflow: (i) generate a set of equiprobable stochastic simulations; (ii) simultaneous stochastic optimization of a mining complex; (iii) select drillhole coordinate with learned policy; (iv) sample drillhole from stochastic simulation; (v) apply EnKF update; and (vi) retrieve informed drilling locations and resulting updated production schedule.

The process begins by inputting a set of stochastic orebody simulations into the simultaneous stochastic optimization framework. These stochastic orebody models represent the uncertain supply of material within the mineral deposit and are simulated using geostatistical techniques (Godoy 2003; Boucher and Dimitrakopoulos 2009). Then, the long-term production schedule is simultaneously optimized by jointly considering the extraction sequence, destination policy and processing stream decisions to maximize net present value and manage technical risk. The optimization considers current information available for orebody modelling which accounts for the uncertainty and local variability of the material grades with a set of stochastic orebody simulations. The formulation for the simultaneous stochastic optimization of mining complexes is detailed in Goodfellow and Dimitrakopoulos (2016; 2017). After obtaining the optimized production schedule given the information prior to infill drilling, a set of features are processed from the long-term production schedule forecasts and the stochastic orebody simulations to define the state of the mining complex. The state is observed by a reinforcement learning agent (or decision maker), which chooses an infill drilling location based on the current neural network policy denoted π_θ. The output action is a continuous output representing the coordinate of the drillhole collar (x, y) and whether to continue drilling. At the corresponding drillhole location, a sample is virtually drilled by retrieving drillhole data from a stochastic simulation of future infill drilling data that is considered here to represent the true deposit. EnKF is then applied to update the stochastic orebody simulations with the additional drilling information retrieved. After the update, the simultaneous stochastic optimization approach is used to reoptimize the production schedule given new information, to determine if the update leads to any changes and improvements in the long-term production schedule and related forecasts. Lastly, the framework is tested several times using a different randomly selected simulation of future drillhole data to ensure similar locations are drilled independent of the realization selected. Details related to each step are discussed in the subsequent sections. In addition, all virtually drilled drillholes are drilled vertically; therefore, the dip and azimuth are not considered and a fixed depth is assumed. These limitations are left to be addressed in future work.

2.2.1 Reinforcement learning background

The reinforcement learning framework applied is an actor-critic reinforcement learning approach (Lillicrap et al. 2015; Sutton and Barto 2018). The agent or decision maker interacts with a mining complex environment E over discrete timesteps to maximize the cumulative reward or return. The infill drilling approach is modelled as a Markov decision process consisting of a state space S, action space A, initial probability distribution with density $p_1(s_1)$, and stationary transition dynamics distribution with conditional density $p(s_{t+1}|s_t, a_t)$ and reward function r. Each timestep t the agent receives an observation of the state of the environment s_t and selects an action a_t. The environment responds by presenting new states s_{t+1} to the agent along with a reward r_t that is generated by the reward function. The agent behaviour is based on a learned policy π_θ that maps states to probabilities of selecting each possible action which can be accomplished by using a neural network policy with a parameter vector θ. The return R_t that is obtained by the agent is the sum of discounted future rewards such that $R_t = \sum_{k=t}^{T} \gamma^{k-t} r\,(s_k, a_k)$ where $\gamma \in (0, 1]$ is the discount factor and T is the terminating timestep. The actor-critic reinforcement learning algorithm is used in this work to determine a policy that maximizes the expected return from the starting state which is denoted by a performance function $J(\pi) = \mathbb{E}[R_1|\pi]$.

Actor-critic reinforcement learning is applied largely due to its success with large continuous action spaces (Lillicrap et al. 2015), which works well for selecting infill drilling locations. In actor-critic reinforcement learning the policy is learned by adjusting the neural network parameters θ in the direction of the performance gradient $\nabla_\theta J(\pi_\theta)$ (Sutton et al. 1999). This increases the probability of selecting actions that led to higher rewards and reduces the likelihood of selecting lower rewards. In actor-critic reinforcement learning, the value function is also learned. The value function is defined as the expected total discounted reward from a given state such that $V^\pi(s) = \mathbb{E}[R_t|s_t = s; \pi]$ and the action value function $Q^\pi(s,a) = \mathbb{E}[R_t|s_t = s, a_t = a;\ \pi]$. In this case the value function represents the incremental improvement achieved by optimizing the long-term production schedule given additional infill drilling data.

Additionally, when neural networks or other nonlinear function approximators are used to approximate a value function it can be unstable or potentially diverge. This is caused for two primary reasons including correlations in the sequence of the observations and the correlations between the action values and target values. These challenges are addressed using a replay buffer and a target network that are periodically updated (Lin 1992; Mnih et al. 2013; Mnih et al. 2015).

2.2.2 Defining the state space

The state of the mining complex environment that is used for selecting infill drilling locations is based on features derived directly from the optimized production schedule, orebody simulations, and available information on site. These provide a numerical representation of the factors that were considered important in this work for drillhole selection. This information can include previously drilled locations, the scheduled period of extraction a drillhole will intersect, and areas of higher or lower grade variability. Each feature is calculated based on the values obtained from the production schedule at timestep t and linear or non-linear transformation of the input parameters denote by the function ϕ:

$$s_t = \phi(\mathrm{sched}_t) \tag{2.1}$$

where the sched_t is the production schedule of the mining complex given the information available at timestep t. These features are the input to the neural network policy π_θ that maps states to actions for infill drillhole selection. Through continuous trial and error, the most beneficial drilling locations are learned by optimizing the parameters of the neural network using a stochastic gradient algorithm (Kingma and Ba 2014). Unlike a stochastic search, context related to the mineral deposit and the mining complex production schedule can be incorporated into the optimization process. This provides insight on which locations to drill within the mining complex based not only on the material uncertainty in the blocks themselves but on the distribution of metal generated downstream at different destinations during each production period. For example, years with higher material variability in terms of metal production may be targeted heavily depending on the impact on the production schedule.

2.2.3 Updating simulations with ensemble Kalman filter (EnKF)

An EnKF framework is applied where the stochastic orebody simulations are updated using new information collected from future infill drilling (Figure 2.2.2). The magnitude of the update is based on the error between the simulated realizations and observed drillhole data that is retrieved from a stochastic simulation of the drillhole data at the location selected. The sampled data represents the drilling data in a real deposit and different randomly selected simulations of the drillhole data are used over several runs to ensure that after each run of the optimization that the drillholes selected are within similar areas. Benndorf (2020) describes the EnKF updating framework for mineral resource applications in further details.

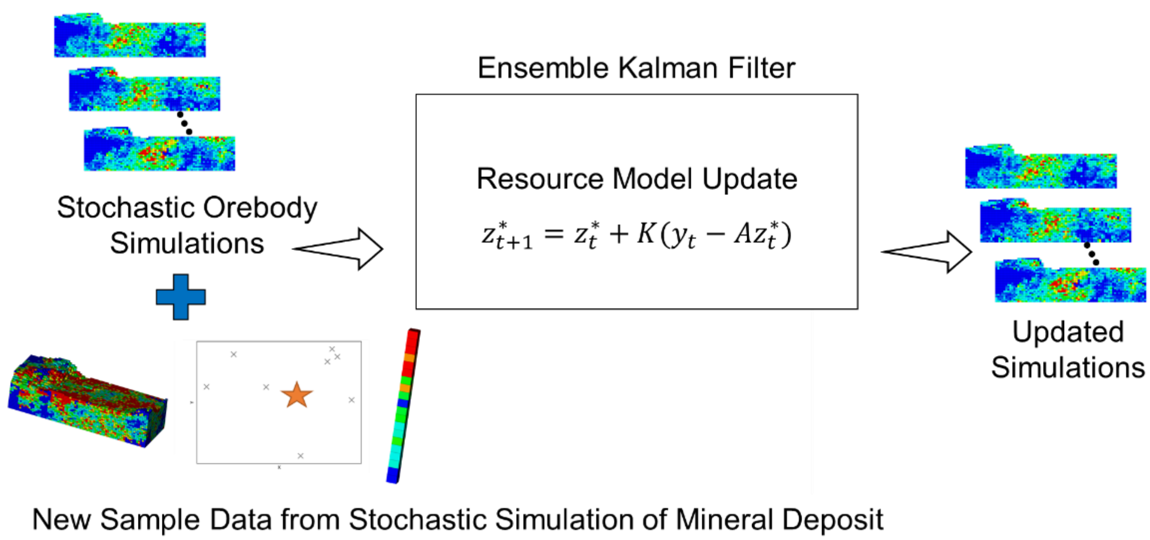


Figure 2.2. EnKF approach for updating stochastic simulations with additional drilling data.

The EnKF framework provides an efficient process for updating a set of simulations $\mathbb{S}$ of a mineral deposit without re-simulating the entire deposit. A spatial random field of the attribute of interest is denoted $\boldsymbol{Z}(\boldsymbol{x})$ where each element $\boldsymbol{Z}(\boldsymbol{x}_j)$ is a random variable associated with a block position at location $\boldsymbol{x}_j$ for $j = \{1, \ldots, M\}$. Each realization of the spatial random field and the corresponding material grade is denoted $\boldsymbol{z}^*(\boldsymbol{x})_{t,\mathscr{s}}$. This provides the predicted value at location $\boldsymbol{x}_j$ given the information gathered up to timestep t and the realization $\mathscr{s}$, where $t = \{1, \ldots, T\}$ and $\mathscr{s} \in \mathbb{S}$. A timestep represents each time new information becomes available by selecting a new drillhole location. The spatial random field is limited to a neighbourhood of blocks in the deposit as it is only desirable to update parts of the realization neighbouring the newly collected additional information. A T by M matrix $\boldsymbol{A}$ represents the contribution of each new drillhole sample up to the present timestep t. The rows of matrix $\boldsymbol{A}$ represent the contribution of each drillhole selected to the posterior simulated orebody realizations. When one of the N potential drilling locations are drilled all the locations intersected by drilling vertically contribute to the update and are included as a one in the matrix $\boldsymbol{A}$. Each block location $\boldsymbol{x}_j$ in the neighbourhood is then updated using the following update:

$$\boldsymbol{z}^*(\boldsymbol{x})_{t+1,\mathscr{s}} = \boldsymbol{z}^*(\boldsymbol{x})_{t,\mathscr{s}} + \underbrace{\boldsymbol{C}_{t,t}\boldsymbol{A}^T\left(\boldsymbol{A}\boldsymbol{C}_{t,t}\boldsymbol{A}^T + \boldsymbol{C}_{v,v}\right)^{-1}}_{\text{Kalman Gain}} \underbrace{\left(\boldsymbol{o}_{t,s} - A\boldsymbol{z}^*(\boldsymbol{x})_{t,\mathscr{s}}\right)}_{\text{Innovation}} \tag{2.2}$$

where $\boldsymbol{z}^*(\boldsymbol{x})_{t+1,s}$ is the updated attribute for each realization s after additional information is collected in timestep t. The covariance matrix $\boldsymbol{C}_{t,t}$ is a N by N matrix approximating the auto- and cross covariance of the random field between each element Z($\boldsymbol{x}_j$). The Kalman gain matrix depicted in Equation 2.2 represents the weighting factor that determines the contribution of the update to those neighbouring locations considered in the update. This includes the covariance matrix $\boldsymbol{C}_{v,v}$ of the measured error in the sample. The second component in the update is the innovation, which represents the error between the predicted outcome of the prior model and the observed drillhole value $\boldsymbol{o}_{t,s} = \boldsymbol{o}_t + \epsilon_{\mathscr{s}}$. The observation $\boldsymbol{o}_{t,s}$ is the data value collected through drilling $\boldsymbol{o}_t$ and a small noise term is added to consider measurement error. The EnKF approach uses the set of simulated scenarios to approximate the covariance matrix and increase computational efficiency. This updating framework is applied to the set of stochastic simulations after a drillhole is sampled.

2.2.4 Actor-critic reinforcement learning

An off-policy actor-critic reinforcement learning agent is applied to optimize infill drilling. The actor-critic architecture is based on the policy gradient theorem (Sutton et al. 1999; Peters et al. 2005). During each timestep t the agent receives an observation from the mining complex $s_t = \phi(\mathrm{sched}_t)$, given the current long-term production schedule (sched_t), and a continuous action $a_t \in \mathbb{R}^3$ is chosen to determine the collar coordinate (northing and easting) and whether drilling should be undertaken. The agent then receives a numerical reward r_{t+1} from the mining complex based on the action taken by running the simultaneous stochastic optimization framework given the new information. The learning framework is illustrated in Figure 2.3.

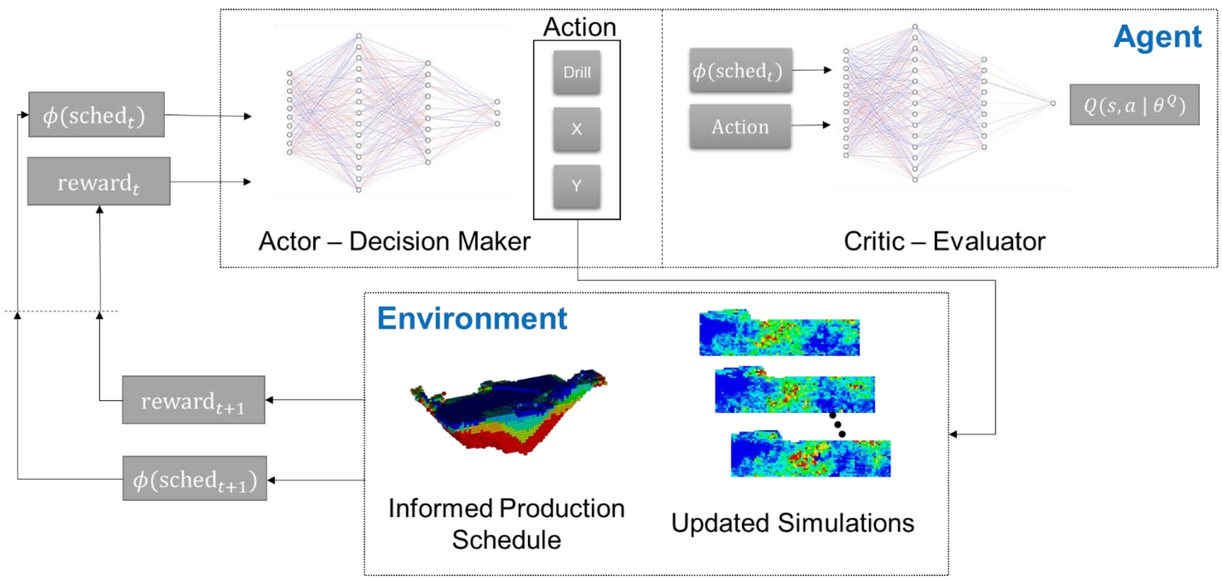


Figure 2.3. Actor-critic reinforcement learning framework for infill drillhole selection.

The actor-critic reinforcement learning algorithm uses an agent composed of two components. The actor learns a parametrized policy that maps the current state of the mining complex s_t to an action $a_t = \mu(s_t|\theta^\mu)$ and controls the decision-making process where θ^μ denotes the parameters of the policy network. Similarly, the critic learns the action-value function $Q(s_t, a_t|\theta^Q)$ through a second set of parameters θ^Q. The action value-function is used to approximate the expected value of taking an action a_t in state s_t parametrized by θ^Q. Upon initialization, the actor $\mu(s_t|\theta^\mu)$ and critic network $Q(s_t, a_t|\theta^Q)$ are randomly initialized with weights θ^μ and θ^Q, respectively. A duplicate of each of these networks are created as a target network to stabilize learning denoted μ' and Q' with initial weights $\theta^{\mu'}$ and $\theta^{Q'}$, respectively. Lastly, a replay buffer denoted $\mathcal{R}$ is initialized. The replay buffer is used to train the network off-policy (Lin 1992). As the agent interacts with the mining complex, each experience it obtains is stored as a tuple (s_t, a_t, r_t, s_{t+1}) in a buffer that contains information from the most recent interaction with the mining complex. These experiences or transitions are randomly sampled to minimize the correlations between samples and update the neural network parameters.

Each episode of learning is a sequence of states, actions, and rewards that end in a terminal state. In the mining complex, the actor decides which locations to infill drill based on the input features and proceeds until the terminal state is reached. The terminal state can occur when the agent decides to no longer continue drilling or a budgetary constraint is exceeded. A random process $\mathcal{N}$ introduces noise to each action to provide adequate exploration of the solution space. This is defined by the exploratory policy $\mu'(s_t) = \mu(s_t|\theta^\mu) + \mathcal{N}$. In addition, the initial state of the environment is received by calculating a set of features $s_1 = \phi(\text{sched}_1)$ derived from the optimized production schedule sched_1 that considers the information available prior to additional drilling. The algorithm proceeds as follows:

1. An action $a_t = \mu(s_t|\theta^\mu) + \mathcal{N}_t$ is selected based on the current policy network and exploration noise.

2. The agent selects the nearest drillhole location to the output coordinates given by the agent's action a_t. If drilling occurs, the drillhole is sampled and used to update a set of simulated realizations $\mathbb{S}_t$ using EnKF. This results in a new set of simulated realizations $\mathbb{S}_{t+1}$ updated to account for the information collected through infill drilling. If the agent decides not to drill step three of the process is skipped and a reward of 0 is received as drilling is not completed.

3. After drilling, a reward r_t is generated by simultaneously optimizing the mining complex given new information. The objective function used for optimizing the long-term production schedule maximizes net present value (NPV) and minimizes deviations from production targets by applying a set of penalties (PEN), see Appendix A. Considering this objective, the reward is computed by determining the difference in the objective function between the optimized

long-term production schedule given new information (sched_{t+1}) and the schedule obtained prior to collecting new information (sched_t) minus the cost of drilling (DC):

$$r_t = \mathbb{E}_{\mathbb{S}_{t+1}}[\mathrm{NPV} - \mathrm{PEN}|\mathrm{sched}_{t+1}] - \mathbb{E}_{\mathbb{S}_{t+1}}[\mathrm{NPV} - \mathrm{PEN}|\mathrm{sched}_t] - \mathrm{DC}\,. \tag{2.3}$$

Note that Equation 2.3 accounts for the most up-to-date information ($\mathbb{S}_{t+1}$) when evaluating the reward. The state s_{t+1} is then reconstructed using the optimized schedule that considers new information and the updated stochastic simulations. The optimized schedule given new information will likely change due to the difference in input and a positive reward is realized if the improvement in the objective function overcomes the cost of additional drilling.

4. Each transition in the environment is stored in a replay buffer $\mathcal{R}$ as a tuple (s_t, a_t, r_t, s_{t+1}) and a batch of P transitions are sampled $(s_i, a_i, r_i, s_{i+1}) \sim \mathcal{R}$ to facilitate learning. The return is calculated using the target network that combines the reward received in timestep i and the predicted action-value provided by the critic target network.

$$y_i = r_i + Q'\left(s_{i+1}, \mu'\left(s_{i+1}\middle|\theta^{\mu'}\right)\middle|\theta^{Q'}\right) \tag{2.4}$$

5. The critic and policy network are updated by minimizing the loss L and using the sampled policy gradient $\nabla_{\theta^{\mu}} J$ to improve the policy, respectively.

$$L = \frac{1}{P}\sum_i \left(y_i - Q(s_i, a_i|\theta^Q)\right)^2 \tag{2.5}$$

$$\nabla_{\theta^{\mu}} J \approx \frac{1}{P}\sum_i \nabla_{\mathrm{a}} Q(s, a|\theta^Q)|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^{\mu}} \mu(s|\theta^{\mu})|_{s=s_i} \tag{2.6}$$

6. Lastly, the target networks are then updated using the updated policy and critic network parameters:

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'} \tag{2.7}$$

$$\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1-\tau)\theta^{\mu'} \tag{2.8}$$

where $\tau \in (0,1]$ is the weighting factor that gradually updates each target network to be closer to current network parameters.

The process is repeated until the agent converges or the maximum number of trials are reached. PyTorch is used for training the neural network (Paszke et al. 2019).

2.2.5 Testing the agent

After learning is completed, the trained actor (neural network policy) is used to select additional infill drilling locations. The stochastic orebody simulations of the deposit prior to additional information are provided as input along with the stochastically optimized production schedule prior to additional information. The random process noise used for action exploration is eliminated, and drillholes are selected until the budget constraint is reached or the trained actor decides it is no longer valuable to continue drilling. The resulting output is the infill drilling configuration that leads to the largest improvement in net present value and/or reduction of technical risk based on the current policy. The learning process is repeated several times with different stochastic simulations of the mineral deposit for sampling drillholes to ensure stability of the results and that similar locations are found independent of the stochastic realization that is sampled.

2.3 Case study in a copper mining complex

The framework developed for selecting additional infill drilling locations is tested in an operating copper mining complex with an open pit mine, several processing streams, stockpiles, and a waste dump facility. The objective of the proposed framework is to determine a set of infill drilling locations that maximize net present value and minimize the risk of deviating from production targets given a fixed drilling budget.

The mining complex considered is shown in Figure 2.4, which includes the allowable material flows from the open pit mine to the customers and the market. A set of stochastic orebody simulations are used to represent the uncertain material supply. The simulated attributes considered include total copper and soluble copper. There are two processing streams with different recoveries. The process plant produces copper concentrate and can recover sulfide and oxide materials. The heap leach destination is used to recover copper from oxide materials and produces a copper cathode product after treatment at a cathode plant. In addition, a stockpile facility is included in the mining complex to allow material to be stored and processed in later periods. The extraction sequence, destination policy, and process stream decisions are optimized with the simultaneous stochastic optimization framework and the resulting production schedule is the initial input to the infill drilling optimization framework.

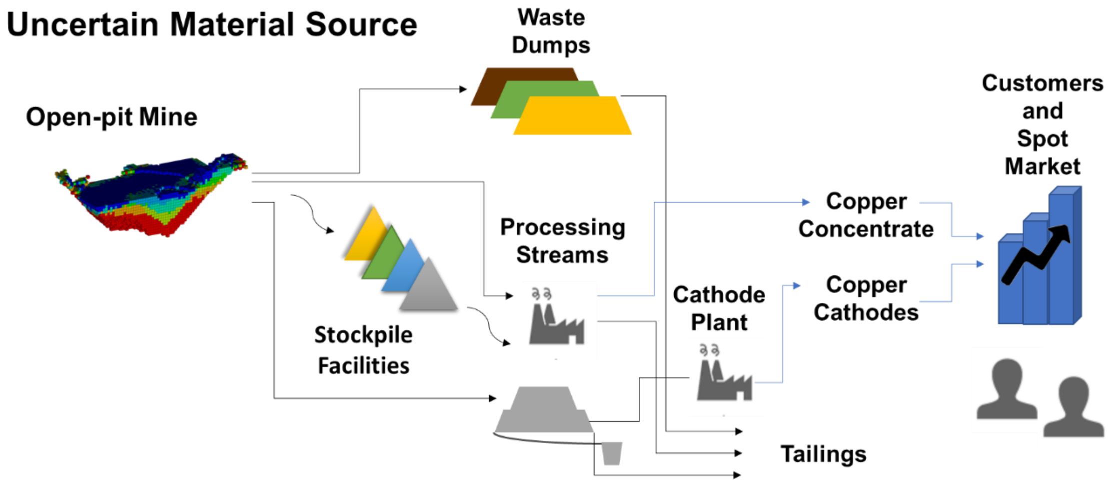


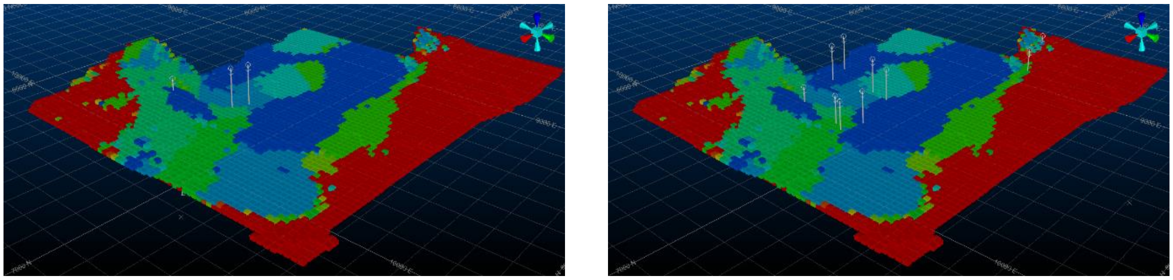
Figure 2.4. The copper mining complex

The case study considers a set of 10 stochastically simulated orebody realizations with a block size of 20 x 20 x 15 m^3 as input and 108,000 blocks. To account for the uncertain mineral deposit, several different geostatistical simulations of future infill drilling data are considered to represent potential samples retrieved from a real mineral deposit. These simulations are different than the stochastic orebody simulations used for optimizing the mining complex. Lastly, there is a budgetary constraint of $1 M and using this budget the algorithm can select from 3640 potential drilling locations.

During training, the actor-critic reinforcement learning agent is left to interact with its environment to learn a set of infill drilling locations that provide the largest improvement in the long-term production schedule. Two example episodes of training the reinforcement learning approach are illustrated in Figure 2.5. The image shows the initial production schedule generated given the information prior to additional infill drilling and the periods of extraction are indicated by the colors of each block. Marked using a circular survey collar are the different drilling locations selected by the actor during each episode are shown. The total reward quantifies the improvement in the objective function obtained through collecting additional information. Additionally, the number of holes drilled in each episode are listed. The number of drillholes selected changes between each episode as the reinforcement learning agent also decides when to stop drilling. The cumulative reward obtained varies significantly depending on where drilling is commenced, and the potential value added by adapting the production schedule to new information. After training is completed by interacting with the mining complex environment, the trained actor (neural network policy) is applied to select the additional infill drilling locations.

The stochastic orebody simulations are provided as input to the proposed framework for infill drilling along with the optimized production schedule and learned reinforcement learning policy (the actor). The random noise is removed from the exploratory policy and the framework is run to select drillholes until either the budgetary constraint is reached or the actor decides it is no longer worthwhile to keep drilling. The infill drilling locations selected with the learned policy are used to forecast the improvement of collecting additional information by evaluating the difference between the production schedule key performance indicators, which are discussed subsequently. Unlike past approaches, the proposed framework seeks infill drilling locations that improve the long-term production schedule forecasts by adapting to new information and do not rely on criteria related to uncertainty of the drilling locations themselves. Infill drilling decisions are selected to improve long-term schedule performs and the mitigate risk related to supply uncertainty within the yearly long-term production schedule.

It should be noted that the stochastic orebody simulations are updated using EnKF when using this approach, which is only optimal if all variables are multivariate Gaussian, the forward observation model $\boldsymbol{A}$ is linear and the random error of the observations are sampled from a Gaussian distribution (Benndorf 2020). Another limitation of applying EnKF is managing the different support sizes between the stochastic orebody model and observed drillhole values (i.e., mining blocks and infill drilling assays), which was not addressed in this work. Lastly, only vertical drillholes were considered which may not be optimal for all orebodies particularly deposits where the boundaries between ore and waste align along a consistent azimuth and dip direction. It may be more advantageous to consider inclined drillholes.



Episode	1
Total Reward	-3.5 x 10^6
Number of Drillholes	3

Episode	40
Total Reward	155.9 x 10^6
Number of Drillholes	10

Figure 2.5. Training examples for different episodes including the number of drillholes drilled and the total cumulative reward obtained.

In Figure 2.6, the resulting production schedule prior to new information (left) is compared to the production schedule with additional information (right). The primary findings reveal that the extents of the final pit are similar between the two production schedules, however, the extraction sequence over the 10-year schedule changes significantly. In the main areas of drilling, marked with a circle, the early extraction sequence decisions change due to additional drilling information by adapting to an area of lower risk to generate further value for the mining complex. As a result of extracting this area earlier in the mine life other areas are pushed back to later periods. This is noticeable in the central part of the deposit as the simultaneous stochastic optimization approach adapts the schedule to manage production targets and increase net present value. In Figure 2.7, the schedule that adapts to additional infill drilling information, which results in a 1.7% increase in the forecasted net present value. Furthermore, the additional information changes the resulting production forecasts leading to a 30 kt increase in recovered copper at the process plant over 10 production years, shown in Figure 2.8. This leads to a significant improvement to the long-term production schedule forecasts which is only obtained by adapting the production schedule to potential new information that is collected via infill drilling. Mining higher grade materials earlier with a higher confidence during the first two production years helps improve the overall production cash flow as drilling provides further confidence for mining these areas caused by taking advantage of the time value of money.

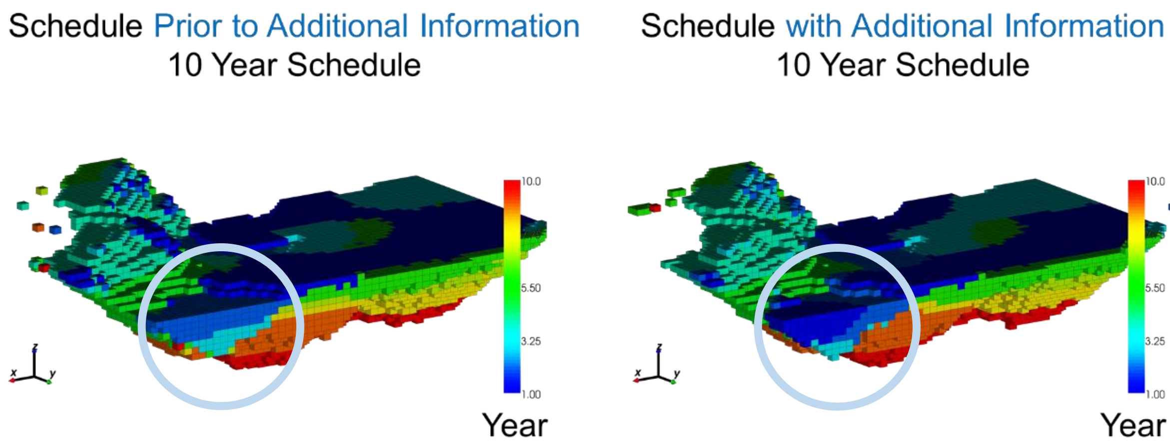


Figure 2.6. Production schedule: (left) prior to additional information; (right) after additional drilling.

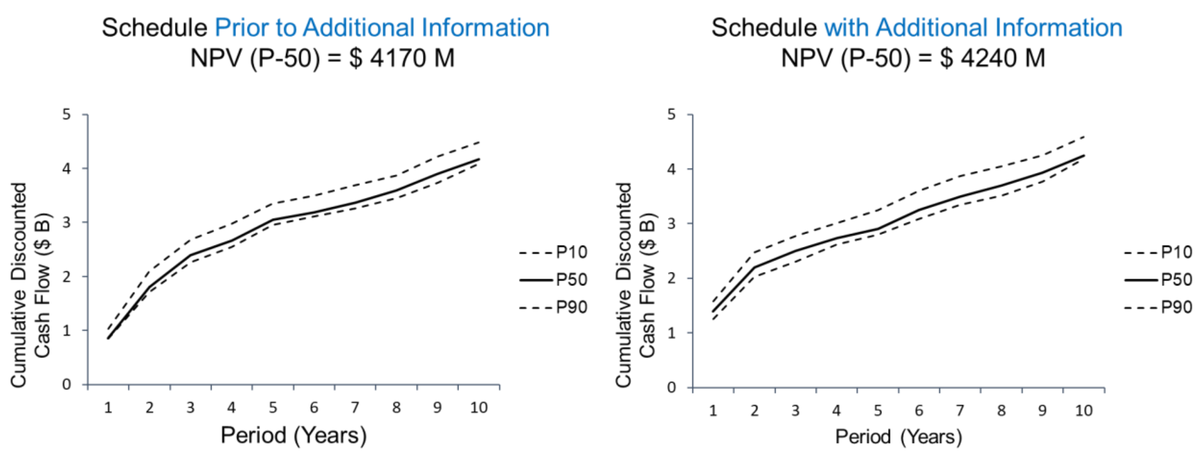


Figure 2.7. Cumulative discounted cash flow given the schedule: (left) prior to additional information; (right) after additional drilling.

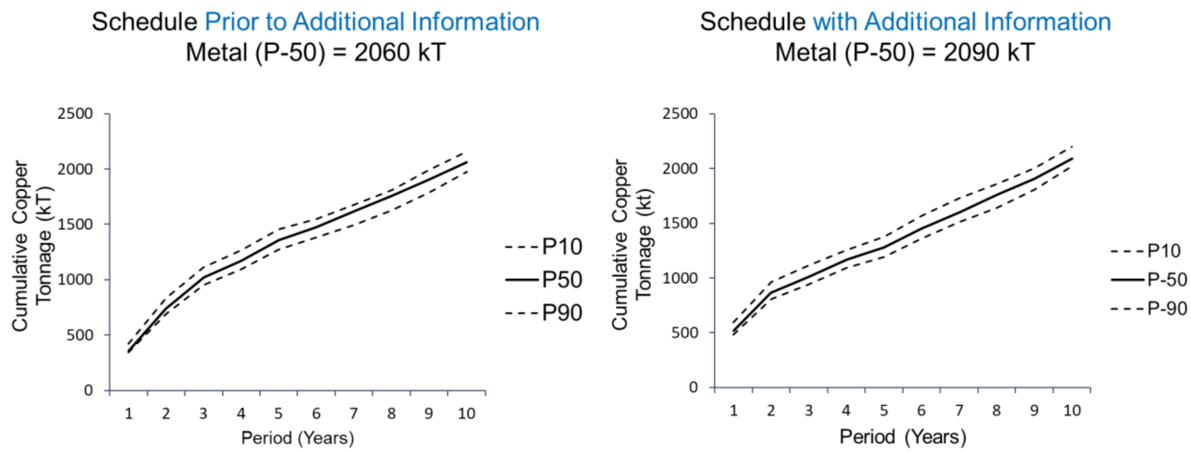


Figure 2.8. Total copper production given the schedule: (left) prior to additional information; (right) after additional drilling.

Five different randomly sampled geostatistical simulations of future infill drilling data are used to represent the potential mineral deposit in five different runs. The proposed framework is run five times where the infill drilling data is retrieved from the respective simulation. The repeated testing of the proposed process using different simulations to represent the mineral deposit shows that similar areas are drilled and identifies a stable area for infill drilling (Figure 2.9). The drillholes selected in each run are shown in Figure 2.9, light grey points show all possible drilling locations in the mining complex and the dark grey region shows the extents of the scheduled areas in the long-term production schedule. The drillholes are highly localized in the northeast portion of the open pit mine and there is considerable overlap in the drillhole locations when using different simulations to represent the sampled infill drilling data. Furthermore, the drilling area is contained within a 250 m x 500 m area. Therefore, the drillhole selection framework identifies a drilling area in a stable way that is not affected by using different or additional realizations.

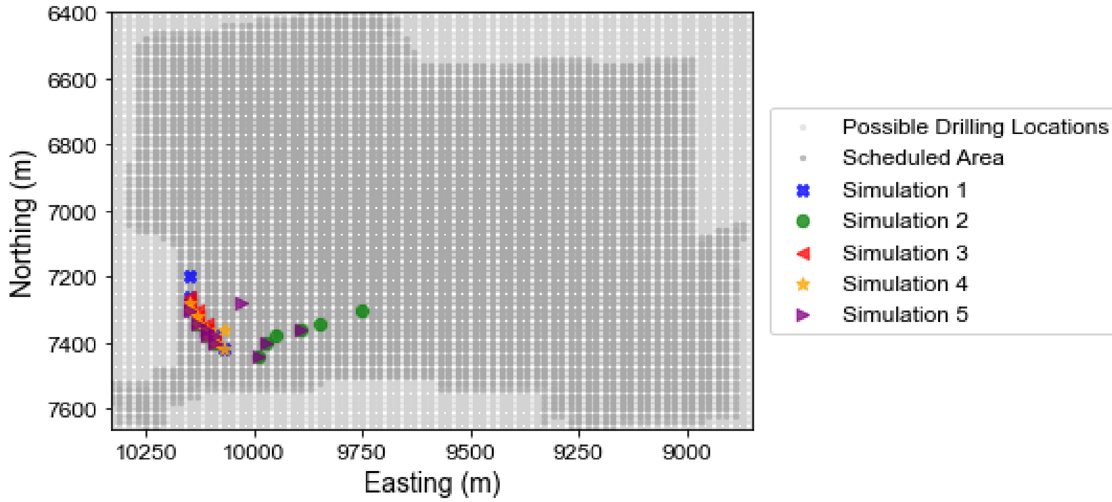


Figure 2.9. The infill drilling locations selected are shown for each run of the algorithm that use a different stochastic simulation to represent the true drillhole data. The light grey shows the possible drilling locations, and the dark grey shows the scheduled areas.

2.4 Conclusions

A new approach has been proposed to determine infill drilling locations in mining complexes that uses actor-critic reinforcement learning. Reinforcement learning guides the framework to valuable drilling locations with a new criterion that links infill drilling to its impact on long-term mine production scheduling decisions. Simultaneous stochastic optimization is applied to quantify the expected improvements to the long-term production schedule based on the new information collected. This provides insight on where drilling should occur to understand the complexity of the mineral deposits by updating the stochastic orebody simulations that are used as inputs for long-term production planning. The proposed approach was tested in a copper mining complex where the results demonstrate significant improvements to the forecasted net present value of the project by adapting the schedule to potential future infill drilling data. This method eliminates the need to use conversion indicators and other proxies to the production schedule value and directly considers the influence of drilling on production scheduling decisions. Actor-critic reinforcement learning is applied to optimize the infill drilling locations and takes advantage of key criteria from the mining complex to select infill drilling locations including considerations related to the uncertainty and local variability of the mineral deposits. Future work should consider allowing the proposed approach to select drillholes with different azimuth and dips to increase the flexibility in the drilling direction. In addition, supplementary information related to the deposit and production schedule could be explored to provide additional context for the drillhole selection process and improve the reinforcement learning agent's ability to distinguish valuable locations to drill.

2.5 Appendix A: Simultaneous stochastic optimization

The simultaneous stochastic optimization framework outlined in Goodfellow and Dimitrakopoulos (2016) uses the following objective function:

$$\max \underbrace{\frac{1}{|\mathbb{S}|}\sum_{i\in\mathcal{S}\cup\mathcal{P}\cup\mathbb{M}}\sum_{t\in\mathbb{T}}\sum_{(h\in\mathbb{H})}\sum_{s\in\mathbb{S}} p_{h,i,t}v_{h,i,t,s}}_{\text{Discounted Revenues and costs}}$$
$$-\underbrace{\frac{1}{|\mathbb{S}|}\sum_{i\in\mathcal{S}\cup\mathcal{P}\cup\mathbb{M}}\sum_{t\in\mathbb{T}}\sum_{(h\in\mathbb{H})}\sum_{s\in\mathbb{S}}\left(c^{+}_{h,i,t}d^{+}_{h,i,t,s}+c^{-}_{h,i,t}d^{-}_{h,i,t,s}\right)}_{\text{Risk discounted penaties for deviations}} \tag{2.9}$$

The optimization objective maximizes the profit accounting for costs and revenues that are incurred to generate the products in the first term and minimizes the penalties from deviations from production targets using the second term. Further details regarding the notation can be found in the reference.

2.6 Appendix B: Hyperparameters

Reinforcement learning frameworks are sensitive to the hyperparameters selected. The top performing parameters are found in Table 2.1 and were obtained by starting with parameters near the published stable baselines for the actor-critic reinforcement learning approach applied (Lillicrap et al. 2015) and exploring nearby parameters with a grid like search. Parameters are assessed based on the highest return achieved during training and the stability of the policy. The discount rate, target network update weight, learning rate and experience batch size were found. The discount rate was used to discriminate between infill drillholes that had subtle differences ensuring the best locations were drilled earlier. A slower update of the target network helped stabilized learning. The Adam optimizer algorithm was used with a learning rate of 1e-5 for the actor and critic network. Mini batches of 32 samples were selected from the replay buffer that had a capacity of 4000 experiences. Additionally, both neural networks are initialized with a single hidden layer containing 256 neurons and apply the ReLU activation function. For exploration noise, a Gaussian process is used with zero mean and a 0.2 standard deviation.

Table 2.1. Hyperparameters for infill drilling

Hyperparameter	**Value**
γ	0.99
τ	0.005
α	1e-5
P	32

2.7 Chapter discussion and next steps

Chapter 2 develops a new stochastic optimization framework for infill drillhole planning, which demonstrates the benefits of optimizing infill drillhole placement given long-term production scheduling criteria. Planning infill drilling by considering its impact on long-term production scheduling decisions is crucial to determine the locations that provide the most information. Past methods do not account for this aspect; thus, the drilling locations may not change the resulting extraction sequence and/or destination policies. When long-term production scheduling decisions are not affected, infill drilling may be an unnecessary investment that could be managed by collecting grade control data during production. The infill drilling framework determines strategic locations to drill that impact the production schedule and should be considered throughout the operational life of a mining complex. Infill drilling information is used to update the stochastic orebody simulations that represent the distribution material grades in the mineral deposits and improve the accuracy of the forecasts obtained using the optimizers developed for short- and long-term production scheduling in the subsequent chapters.

3 Simultaneous stochastic optimization of an open-pit mining complex with preconcentration using reinforcement learning

A short-term simultaneous stochastic optimization framework for mining complexes is developed in this chapter. The framework extends the simultaneous stochastic optimization framework by integrating preconcentration facilities into the optimization formulation along the other critical components of a mining complex. In addition to supply uncertainty related to the material grades, a response rank parameter is stochastically simulated to account for the grade-by-size behaviour of the extracted materials. Certain materials with a beneficial grade-by-size response can be preconcentrated via screening to eliminate coarse waste material from the processing material feed, which eliminates unnecessary crushing and grinding of coarse waste material. The benefits of preconcentration facilities are evaluated in a case study at large-scale copper mining complex. In addition, actor-critic reinforcement learning is applied and combined with stochastic programming techniques to optimize the short-term production schedule by learning an adaptive policy for optimizing the short-term production schedule.

3.1 Introduction

A mining complex is a multifaceted engineering system where valuable minerals are extracted, transported, processed and delivered to customers and the market (Whittle 2007; Pimentel et al. 2010; Goodfellow and Dimitrakopoulos 2016; Montiel and Dimitrakopoulos 2018). Mining operations are planned to leverage the extracted material properties and support the optimized interconnectivity between components. Materials are supplied from open-pit and/or underground mines, then transported downstream to stockpiles, preconcentration facilities, waste dumps or one of several processing facilities to generate marketable products (Hustrulid et al. 2013). Stockpiles store materials for future use. Preconcentration facilities separate and sort materials prior to processing; for example, screening and bulk sorting facilities liberate valuable products from waste based on distinguishable material properties (Bowman and Bearman 2014). Waste dumps are areas for disposing of material without value and processing destinations (i.e., process plants and leach pads) recover products to be sold to the market.

An optimized short-term production schedule is required to satisfy the expected long-term yearly forecasts in an operating mining complex (Smith 1998). Short-term production periods can be days, weeks or months depending on the planning timeframe addressed. Advanced simultaneous stochastic optimization approaches model a mining complex with a stochastic mathematical program and apply metaheuristic solving techniques to jointly optimize several short-term scheduling aspects (Both and Dimitrakopoulos 2020; 2021). This includes the extraction sequence, destination policy, stockpiling, equipment allocation and processing stream decisions. The main advantages of simultaneous stochastic optimization are: (1) all scheduling decisions are optimized jointly (Goodfellow and Dimitrakopoulos 2016); and (2) a set of simulated realizations are used to represent the uncertainty and variability of the mineral deposit (Goovaerts 1997). Significantly higher project value is achieved by simultaneously optimizing all components and exploiting advantageous synergies (Levinson and Dimitrakopoulos 2020b; Saliba and Dimitrakopoulos 2020). In contrast, conventional mine planning approaches consider a subset of critical decisions optimized independently over several steps and ignore uncertainty in the quality and quantity of material (Hoerger et al. 1999b; Whittle and Burks 2010). Instead of stochastic simulations, a single estimated (average type) block model representing the mineral deposit is used to quantify the quality and quantity of materials in each block (Fytas et al. 1993). Conventional approaches often aggregate mining blocks into larger units, failing to represent the material selectivity of an operating mine, and apply decomposition methods to reduce the complexity of the optimization formulation (Blom et al. 2016; Blom et al. 2018; 2019; Kozan and Liu 2016; Upadhyay and Askari-Nasab 2019). Each simplification reduces the possibility of finding additional value. Concerns regarding the use of estimated models of a mineral deposit have been addressed for decades (Ravenscroft 1992; Dowd 1994; Dimitrakopoulos et al. 2002) and led to the development of stochastic mathematical programming approaches to manage material uncertainty directly in the optimization formulation (Dimitrakopoulos and Ramazan 2008; Ramazan and Dimitrakopoulos 2013). Despite considering several parts of a mining complex and managing technical risk, simultaneous stochastic optimization for short-term production scheduling frameworks have not included preconcentration facilities that contribute significantly to processing efficiency and productivity. This paper introduces a new stochastic programming formulation for simultaneously optimizing short-term production schedules with preconcentration. For reasons discussed subsequently, reinforcement learning and stochastic mathematical programming are combined to solve the corresponding large-scale stochastic optimization formulation.

Preconcentration is integrated into the proposed stochastic programming formulation for short-term production scheduling, accounting for a major mining complex component with considerable upside potential (Burns and Grimes 1986). Using sorting and separating technologies, preconcentration facilities can liberate valuable products from waste at a coarser scale then traditional processing methods (Bowman and Bearman 2014; Espejel et al. 2017; Fathollahzadeh et al. 2021); therefore, reducing the quantity of waste material that unnecessarily enters the energy intensive crushing and grinding circuit in a process plant. This is accomplished by leveraging the grade-by-size behaviour of material, which relates to certain materials having higher metal concentration in specific size fractions after breakage (Carrasco et al. 2016a). Fathollahzadeh et al. (2021) recently consider jointly optimizing long-term preconcentration decisions, however, material uncertainty and short-term considerations are ignored. A novel contribution of this work is preconcentration decisions are simultaneously optimized with the short-term production schedule, while accounting for material uncertainty. The uncertain material grade-by-size behaviour at a screening preconcentration facility is modeled with a simulated response rank parameter and a coarse liberation model, defined in Carrasco et al. (2016a). Preconcentration facilities are even more critical for operations in the short-term since these technologies improve processing efficiency, by providing added selectivity for certain materials. This alongside the environmental benefits of preconcentration make it desirable (Adair et al. 2020).

Current algorithms for solving large-scale production scheduling formulations in mining typically rely on meta-, hyper- or math-heuristics (Goodfellow and Dimitrakopoulos 2016; Lamghari and Dimitrakopoulos 2020; Lamghari et al. 2022). However, these methods can be inefficient as the heuristics are static and rely on predetermined strategies for adjusting decisions. These strategies do not learn from the problem or improve over time. Reinforcement learning is another approach for solving stochastic mathematical programming formulations of large-scale problems (Powell 2022). Reinforcement learning algorithms can be used to learn a policy to control heuristics based on contextual information related to the problem and a reward generated by the reinforcement learning environment. This provides an enhanced heuristic technique that adapts to the problem at hand.

In this work, the proposed framework combines reinforcement learning with stochastic programming approaches to optimize the short-term production schedule in real-world mining complexes. The reinforcement learning agent drives the optimization process by learning a heuristic for selecting new extraction sequences with a parametrized policy. The policy is improved by learning behaviors that maximize the expected future rewards (or return). During each iteration of the optimization framework, the agent proposes an extraction sequence using the learned policy and the remaining downstream decisions are optimized with exact methods and a commercial solver, in this case CPLEX (IBM 2017). Because the downstream decisions are optimal, this provides stable rewards for the agent, which accurately guides the agent towards the best production schedule. Additionally, exact solving methods are used to efficiently solve certain parts of the stochastic programming formulation when the decisions of the learned heuristics are fixed.

Efforts to combine stochastic mathematical programming with reinforcement learning to simultaneously optimize a mining complex have received little attention to date. The latest developments that apply reinforcement learning in mining include adapting the destination policy – the location of where to send extracted materials – under various operating conditions (Paduraru and Dimitrakopoulos 2018; 2019) and generating adaptive policies that modify the short-term production schedule including the extraction sequence and destination policy under equipment and grade uncertainty (Kumar et al. 2020; Kumar and Dimitrakopoulos 2021). The first set of approaches only consider learning an adaptive destination policy and do not consider extraction sequence decisions or preconcentration (Paduraru and Dimitrakopoulos 2018; 2019). While the second set of approaches disregard critical operational aspects including stockpiles, preconcentration facilities and operational access requirements for the extraction sequence (Kumar et al. 2020; Kumar and Dimitrakopoulos 2021). These limitations are addressed in this work. Furthermore, fleet dispatching systems for minimizing greenhouse gas emissions using multi-agent Q-learning have been investigated, however, this approach only considers equipment routing (Huo et al. 2023). Although, each reinforcement learning application suffers from its own limitations, the studies demonstrate that an effectively designed reinforcement learning approach can learn a policy to adapt production scheduling decisions and enhance baseline performance.

This paper presents two contributions to the simultaneous stochastic optimization of short-term production schedules. First, a new short-term stochastic programming formulation is proposed that integrates preconcentration, stockpiling, multiple processing destinations and material uncertainty. Second, a novel simultaneous optimization framework is developed that combines reinforcement learning and exact solving methods to solve the short-term production schedule. The combination of reinforcement learning, and stochastic mathematical programming is expected to be an advantageous method for optimizing the short-term production schedule.

The remaining components of this paper proceed as follows. First, the problem statement is detailed. Then, a new stochastic mathematical programming formulation is proposed for simultaneously optimizing the short-term production schedule with preconcentration. A novel solution method is then outlined for solving the stochastic programming formulation by combining reinforcement learning with exact solving methods. Finally, the stochastic programming model and proposed solution approach is tested in a large-scale operating mining complex. Conclusions and future work follow.

3.2 Problem Statement

A mining complex is defined here as a of set of open-pit mines ($\mathcal{M}$), stockpiles ($\mathcal{S}$), preconcentration ($\mathcal{F}$), processing ($\mathcal{P}$) and waste ($\mathcal{W}$) facilities, as shown in Figure 3.1. Materials at the mines are clustered into a set of groups ($\mathcal{G}$) with similar response rank (RR) and material properties related to the metal grades (Q). The material routes in a mining complex are modeled as a directed graph where nodes define the different locations in a mining complex and directed arcs represent the allowable material routes. Sources of extracted material $\{\mathcal{M} \cup \mathcal{S} \cup \mathcal{F}\}$ are sent to downstream destinations $\{\mathcal{S} \cup \mathcal{F} \cup \mathcal{P} \cup \mathcal{W}\}$ and depending on the material properties are transformed into valuable products. For each mineral deposit, the volume of material to be considered for mining are represented by a three-dimensional block model. A set of stochastic simulations (S) represent the uncertain quality and quantity of material within each block in the block model, as seen in Figure 3.2. The equiprobable stochastic simulations reproduce the statistics of the sample data and are obtained with geostatistical techniques (Goovaerts 1997; Boucher and Dimitrakopoulos 2009).

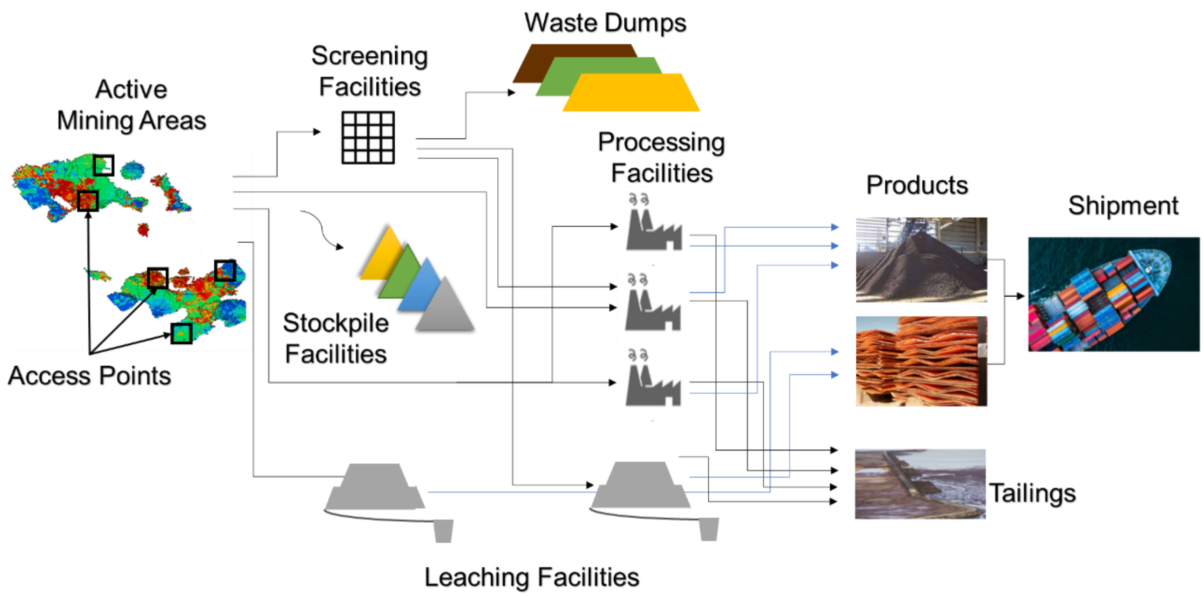


Figure 3.1. An example of a copper mining complex. On the left, the squares overlaying the active mining areas represent the mining access points.

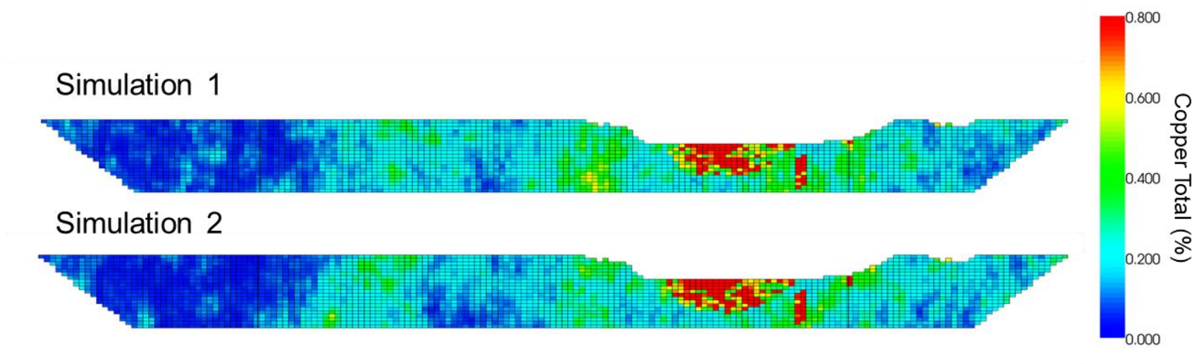


Figure 3.2. An example of two stochastically simulated orebody (block) models, cross-sections of the mineral deposit, representing the uncertain total copper grade in the material to be mined.

Within the mines, there are several mining access points (AP) for extracting materials. These are shown on the left side of Figure 3.1. The short-term production schedule is guided by access point availability because they provide haulage infrastructure with ramps and benches to effectively load and haul the surrounding materials. A high precision GPS is used to determine the geographical position (POS) of each access point. In an operating mine, a short-term mine planner must determine the appropriate extraction sequence for mining. This includes the direction and quantity of material to mine from each access point to progress the mining sequence. This is often a manual and iterative process. An extraction sequence is desired that balances the capacities at all the locations in a mining complex and satisfies long-term targets, without compromising value. Typically, mining operations will determine the short-term production schedule by considering the material properties in the mining blocks surrounding each access point to determine the extraction sequence, while also evaluating the preferred direction of the long-term plan. These properties provide the decision maker with an understanding of the downstream behaviour of material at stockpiles, preconcentration and processing facilities. Then, the remaining downstream components are optimized independently.

As discussed above, preconcentration provides major benefits for mining operations as it eliminates waste material from entering processing facilities based on distinguishable material properties. Screening preconcentration facilities are configured herein to separate material into coarse and fine material streams based on a targeted material size. Those materials with beneficial grade-by-size behaviour are separated at the screen. An uncertain response rank (RR) parameter is used to quantify the grade-by-size behaviour of material at a screen (Carrasco et al. 2016a; Carrasco et al. 2016b). Samples are collected to determine the response rank by running industrial scale screening equipment pilots and laboratory testing methods (Carrasco et al. 2016a). The percentage of material passing a screen into the fine feed at a preconcentration facility is defined by the mass response ($\mathrm{MR} \in (0,1]$) and the remaining material is diverted into the coarse material feed. Using the simulated RR, a response factor (RF) is calculated to reflect the expected upgrade in material quality for those materials passing the screen using an empirical coarse liberation model:

$$\mathrm{RF} = \exp\left(-\mathrm{RR} * \frac{\log(\mathrm{MR})}{200}\right)$$

where the benefits vary for different response ranks and mass responses (Carrasco et al. 2016a; Fathollahzadeh et al. 2021). Figure 3.3 shows that higher response ranks are more beneficial than lower responses and higher response factors can be obtained at the expense of less material passing a screen, if grade-by-size behaviour exists (i.e., RR $\neq 0$). The coarse liberation model is used to simulate the separation behaviour of materials after they are sent to preconcentration facilities.

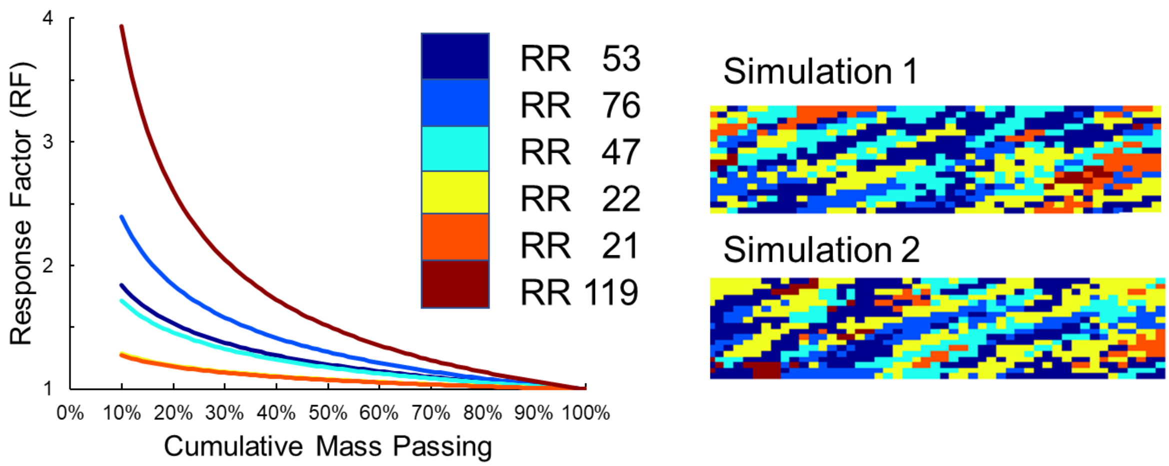


Figure 3.3. On the left, the response factor vs mass passing (response) for different response ranks is shown. On the right, a cross-section of the simulated response rank in two different realizations are shown.

The optimized short-term production schedule must provide a mineable extraction sequence, destination policy, stockpiling and preconcentration decisions to help operations meet long-term planning forecasts and mitigate risk. In the following section, the stochastic mathematical program for optimizing the short-term production schedule with preconcentration is outlined.

3.3 Stochastic mathematical programming formulation for a mining complex with preconcentration

The proposed stochastic programming formulation for optimizing a mining complex with preconcentration via screening is an extension to Goodfellow and Dimitrakopoulos (2016).The formulation optimizes several major short-term production scheduling decisions to create an operative plan which includes determining the extraction sequence, destination policy, stockpiling and preconcentration decisions. The major extensions to this model are that preconcentration decisions are included and the behaviour of screening facilities is modeled with a coarse liberation model to account for the material response rank, mass response and response factor. These considerations impact the quality and quantity of coarse and fine material at each location in the mining complex. Definitions and notations are now introduced.

Sets:

T	Set of short-term production scheduling time periods, indexed by t.
$\mathcal{M}$	Set of mines.
$\mathcal{S}$	Set of stockpiles.
$\mathcal{F}$	Set of preconcentration facilities.
$\mathcal{P}$	Set of processing facilities.
$\mathcal{W}$	Set of waste facilities.
$\mathcal{D}$	Set of destinations $\mathcal{D} = \{\mathcal{S} \cup \mathcal{F} \cup \mathcal{P} \cup \mathcal{W}\}$.
$\mathcal{G}$	Set of material groups, indexed by g.
$\mathcal{I}(i)$	Set of locations that supply material to location $i \in \mathcal{D}$.
$\mathcal{O}(i)$	Set of locations that receive material from location $i \in \mathcal{D}$.
B_m	Set of blocks in the block model for each mine $m \in \mathcal{M}$, indexed by b.
S	Set of stochastic simulations of the uncertain material properties, indexed by s.
Q	Set of material quality properties considered in the optimization model, indexed by q.

Parameters:

$w_{b,s}$	Mass of block $b \in B_m$ in scenario $s \in S$.
$g_{b,q,s}$	Grade quality attribute $q \in Q$ in block $b \in B_m$ in scenario $s \in S$.
$\mathrm{RF}_{b,q,i,s}$	Response factor for attribute $q \in Q$ in block $b \in B_m$ at preconcentration facility $i \in \mathcal{F}$ in scenario $s \in S$.

$\mathrm{MR}_{b,i,s}$	Mass response factor for block $b \in B_m$ at preconcentration facility $i \in \mathcal{F}$ in scenario $s \in S$.
$\theta_{b,g,s}$	Group membership parameter. Takes a value of one if block $b \in B_m$ is in group $g \in \mathcal{G}$ in scenario $s \in S$; zero otherwise.
$p_{q,i,t}$	Selling price per unit of product produced with attribute $q \in Q$ and recovered at location $i \in \mathcal{P}$ in period $t \in T$.
$c_{i,t}$	Unit cost to extract, handle and process material at location $i \in \{\mathcal{M} \cup \mathcal{S} \cup \mathcal{P} \cup \mathcal{W}\}$ in period $t \in T$.
$\mathrm{sc}_{i,t}$	Unit cost for preconcentrating material via screening at location $i \in \mathcal{F}$ in period $t \in T$.
$r_{q,i,t}$	Recovery of property $q \in Q$ at processing location $i \in \mathcal{P}$ and period $t \in T$.
$\mathrm{pc}^{-}_{i,t}, \mathrm{pc}^{+}_{i,t}$	Penalty cost for a positive (+) or negative (-) deviation from short-term capacity at location $i \in \{\mathcal{M} \cup \mathcal{D}\}$ in period $t \in T$.
$\mathrm{pc}^{\mathrm{TT},-}$	Penalty cost for negative deviations from long-term tonnage target TT.
$\mathrm{pc}^{\mathrm{TP},-}_{q,i}$	Penalty cost for negative deviations from long-term total of property $q \in Q$ processed at location $i \in P$.
$L_{i,t}$	Lower bound on short term capacity $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T$.
$U_{i,t}$	Upper bound on short term capacity $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T$.
$\mathrm{TP}_{q,i}$	Forecasted long-term metal production for property $q \in Q$ at processing facility $i \in \mathcal{P}$.
TT	Forecasted long-term total tonnage mined over all access points and mines.

Decision variables:

$x_{b,t}$	Binary variable for each block $b \in B_m$ and production period $t \in T$. Takes a value of one if block b is mined in period t; zero otherwise.

$\kappa_{g,j,t}$ — Binary variable determines whether material mined in group $g \in G$ is sent to destination $j \in \mathcal{D}$ during period $t \in T$. Takes a value of one if group g is sent to destination j in period t; zero otherwise.

$z_{i,j,t}^{\text{coarse}}, z_{i,j,t}^{\text{fine}}$ — Binary decision variable defining the destination of material separated at screening facility $i \in \mathcal{F}$ and sent to destination $j \in \mathcal{O}(i)$ in period $t \in T$. Takes a value of one if coarse or fine material is sent from screen i to destination j during periot t; zero otherwise.

$y_{i,j,t}$ — Continuous variable for determining the fraction of material sent from location $i \in \mathcal{S}$ to destination $j \in \mathcal{O}(i)$ in period $t \in T$.

$\delta_{i,t,s}^{+}, \delta_{i,t,s}^{-}$ — Recourse variables measures positive (+) and negative (-) deviations from short-term production targets at location $i \in \mathcal{M} \cup \mathcal{D}$ during period $t \in T$ and under scenario $s \in S$.

$\delta_{s}^{\text{TT},-}$ — Recourse variable measures negative deviations from long-term total tonnage (TT) targets under scenario $s \in S$, positive deviations are not considered.

$\delta_{q,i,s}^{\text{TP},-}$ — Recourse variable measures negative deviations from long-term total processed (TP) targets for property $q \in Q$ at processing destination $i \in \mathcal{P}$ under scenario $s \in S$, positive deviations are not considered.

$\lambda_{g,i,t,s}^{\text{coarse}}$ — Continuous variable defining the total mass of coarse material extracted in group $g \in \mathcal{G}$ to be sent to location $i \in \mathcal{D}$ in period $t \in T$ and scenario $s \in S$.

$\lambda_{g,i,t,s}^{\text{fine}}$ — Continuous variable defining the total mass of fine material extracted in group $g \in \mathcal{G}$ to be sent to location $i \in \mathcal{D}$ in period $t \in T$ and scenario $s \in S$.

$\gamma_{q,g,i,t,s}^{\text{coarse}}$ Continuous variable defining the total mass of coarse material property $q \in Q$ extracted from group $g \in G$ to be sent to location $i \in \mathcal{D}$ in period $t \in T$ and scenario $s \in S$.

$\gamma_{q,g,i,t,s}^{\text{fine}}$ Continuous variable defining the total mass of fine material property $q \in Q$ extracted from group $g \in G$ to be sent to location $i \in \mathcal{D}$ in period $t \in T$ and scenario $s \in S$.

$\text{MASS}_{i,t,s}^{\text{coarse}}$ Continuous variable denoting the quantity of coarse material received at location $i \in \{\mathcal{F} \cup \mathcal{P} \cup \mathcal{W}\}$ in period $t \in T$ and scenario $s \in S$.

$\text{MASS}_{i,t,s}^{\text{fine}}$ Continuous variable denoting the quantity of fine material received at location $i \in \{\mathcal{F} \cup \mathcal{P} \cup \mathcal{W}\}$ in period $t \in T$ and scenario $s \in S$.

$\text{STK}_{i,t,s}$ Continuous variable denoting the mass of material in the stockpile location $i \in \mathcal{S}$ in period $t \in T$ and scenario $s \in S$ where $\text{STK}_{i,0,s}$ is initialized to the uncertain existing stockpile quantities. Only coarse materials are stored at stockpiles.

$\text{METAL}_{q.i,t,s}^{\text{coarse}}$ Continuous variable denoting the amount of coarse material property $q \in Q$ received at location $i \in \{\mathcal{F} \cup \mathcal{P} \cup \mathcal{W}\}$ in period $t \in T$ and scenario $s \in S$.

$\text{METAL}_{q,i.t,s}^{\text{fine}}$ Continuous variable denoting the amount of fine material property $q \in Q$ received at location $i \in \{\mathcal{F} \cup \mathcal{P} \cup \mathcal{W}\}$ in period $t \in T$ and scenario $s \in S$.

$\text{MSTK}_{q,i,t,s}$ The quantity of material property $q \in Q$ at each stockpiling location $i \in \mathcal{S}$ in period $t \in T$ and scenario $s \in S$; $\text{MSTK}_{q,i,0,s}$ is initialized to the uncertain existing stockpile quantities.

Objective function:

The objective function in Eq. 3.1 is composed of several parts: part 1 maximizes the revenues obtaining from processing the coarse and fine material and generating valuable products at the processing facilities; part 2 minimizes costs of mining, stockpiling and processing to produce valuable products; part 3 minimizes penalties for positively and negatively deviating from short-term capacity constraints; part 4 minimizes the cost of using preconcentration facilities; and part 5 minimizes the cost of deviating from long-term production forecasts to align schedules. The main contributions of this model in the objective function are the newly added parts 4 and 5 of the objective. In part 4, preconcentration costs are accounted for in the optimization model allowing for the trade-off between processing all material together or separating material into coarse and fine fractions to be evaluated. Part 5 links short-term scheduling to long-term targets by penalizing the objective function for misalignment between the schedules of different timescales.

$$\max \frac{1}{\|S\|}\left(\underbrace{\sum_{s\in S}\sum_{t\in T}\sum_{i\in\mathcal{P}}\sum_{q\in Q}\mathrm{p}_{q,t}r_{q,i,t}\left(\mathrm{METAL}_{q,i,t,s}^{\mathrm{coarse}}+\mathrm{METAL}_{q,i,t,s}^{\mathrm{fine}}\right)}_{\text{Part 1}}\right.$$
$$\underbrace{-\sum_{s\in S}\sum_{t\in T}\sum_{i\in\{\mathrm{M}\cup\mathcal{D}\setminus\mathcal{F}\}}\mathrm{c}_{i,t}\left(\mathrm{MASS}_{i,t,s}^{\mathrm{coarse}}+\mathrm{MASS}_{i,t,s}^{\mathrm{fine}}\right)}_{\text{Part 2}}\underbrace{-\sum_{s\in S}\sum_{t\in T}\sum_{i\in\mathcal{M}\cup\mathcal{D}}\left(\mathrm{pc}_{i,t}^{-}\delta_{i,t,s}^{-}+\mathrm{pc}_{i,t}^{+}\delta_{i,t,s}^{+}\right)}_{\text{Part 3}}$$
$$\left.\underbrace{-\sum_{s\in S}\sum_{t\in T}\sum_{i\in\mathcal{F}}\mathrm{sc}_{i,t}\left(\mathrm{MASS}_{i,t,s}^{\mathrm{coarse}}+\mathrm{MASS}_{i,t,s}^{\mathrm{fine}}\right)}_{\text{Part 4}}\underbrace{-\sum_{s\in S}\left(\mathrm{pc}^{\mathrm{TT},-}\delta_{s}^{\mathrm{TT},-}+\sum_{q\in Q}\sum_{i\in\mathcal{P}}\mathrm{pc}_{q,i}^{\mathrm{TP},-}\,\delta_{\mathrm{q,i,s}}^{\mathrm{TP},-}\right)}_{\text{Part 5}}\right) \tag{3.1}$$

Subject to:

Preconcentration constraints ensure material separated into coarse and fine materials are only sent to a single destination:

$$\sum_{j\in\mathcal{O}(i)} z_{i,j,t}^{\mathrm{coarse}} = 1 \quad \forall i \in \mathcal{F}, t \in T. \tag{3.2}$$

$$\sum_{j\in\mathcal{O}(i)} z_{i,j,t}^{\mathrm{fine}} = 1 \quad \forall i \in \mathcal{F}, t \in T. \tag{3.3}$$

Mined material constraints calculate the quantity and quality of coarse and fine material in each group that can be sent to each allowable destination. These constraints account for the mass response (MR) and response factor (RF) and divide the materials and their qualities into coarse and fine materials based on their behaviour at the screening facility:

$$\lambda_{g,i,t,s}^{\text{coarse}} = \sum_{m\in\mathcal{M}} \sum_{b\in B_m} \text{MR}_{b,i,s}\theta_{b,g,s}w_{b,s}x_{b,t} \quad \forall g \in G, i \in \mathcal{D}, t \in T, s \in S. \tag{3.4}$$

$$\lambda_{g,i,t,s}^{\text{fine}} = \sum_{m\in\mathcal{M}} \sum_{b\in B_m} (1 - \text{MR}_{b,i,s})\theta_{b,g,s}w_{b,s}x_{b,t} \quad \forall g \in G, i \in \mathcal{D}, t \in T, s \in S. \tag{3.5}$$

$$\begin{aligned}\gamma_{q,g,i,t,s}^{\text{coarse}} = \sum_{m\in\mathcal{M}} \sum_{b\in B_m} (1 - \text{MR}_{b,i,s}\text{RF}_{b,q,s})\theta_{b,g,s}w_{b,s}g_{b,q,s}x_{b,t} \\ \forall q \in Q, g \in G, i \in \mathcal{D}, t \in T, s \in S.\end{aligned} \tag{3.6}$$

$$\begin{aligned}\gamma_{q,g,i,t,s}^{\text{fine}} = \sum_{m\in\mathcal{M}} \sum_{b\in B_m} \text{MR}_{b,i,s}\theta_{b,g,s}w_{b,s}g_{b,q,s}\text{RF}_{b,q,s}x_{b,t} \\ \forall q \in Q, g \in G, i \in \mathcal{D}, t \in T, s \in S.\end{aligned} \tag{3.7}$$

Material balancing constraints calculate the amount of coarse and fine material and quantity of metal from the given material properties at each destination in the mining complex:

$$\begin{aligned}\text{MASS}_{i,t,s}^{\text{coarse}} = \sum_{g\in G} \kappa_{g,i,t}\lambda_{g,i,t,s}^{\text{coarse}} + \sum_{j\in\mathcal{I}(i)\cap\mathcal{F}} z_{j,i,t}^{\text{coarse}}\text{MASS}_{j,t,s}^{\text{coarse}} \\ + \sum_{j\in\mathcal{I}(i)\cap\mathcal{S}} y_{j,i,t}\text{STK}_{j,t,s} \quad \forall i \in \mathcal{D}\backslash\mathcal{S}, t \in T, s \in S\,.\end{aligned} \tag{3.8}$$

$$\text{MASS}_{i,t,s}^{\text{fine}} = \sum_{g\in G} \kappa_{g,i,t}\lambda_{g,i,t,s}^{\text{fine}} + \sum_{j\in\mathcal{I}(i)\cap\mathcal{F}} z_{j,i,t}^{\text{fine}}\text{MASS}_{j,t,s}^{\text{fine}} \quad \forall i \in \mathcal{D}\backslash\mathcal{S}, t \in T, s \in S. \tag{3.9}$$

$$\begin{aligned}\text{METAL}_{q,i.t,s}^{\text{coarse}} = \sum_{g\in G} \kappa_{g,i,t}\gamma_{q,g,i,t,s}^{\text{coarse}} + \sum_{j\in\mathcal{I}(i)\cap\mathcal{F}} z_{j,i,t}^{\text{coarse}}\text{METAL}_{j,t,s}^{\text{coarse}} \\ + \sum_{j\in\mathcal{I}(i)\cap\mathcal{S}} y_{j,i,t}\text{MSTK}_{q,j,t,s} \quad \forall q \in Q, i \in \mathcal{D}\backslash\mathcal{S}, t \in T, s \in S.\end{aligned} \tag{3.10}$$

$$\begin{aligned}\text{METAL}_{q,i,t,s}^{\text{fine}} = \sum_{g\in G} \kappa_{g,i,t}\gamma_{q,g,i,t,s}^{\text{fine}} + \sum_{j\in\mathcal{I}(i)\cap\mathcal{F}} z_{j,i,t}^{\text{fine}}\text{METAL}_{j,t,s}^{\text{fine}} \\ \forall q \in Q, i \in \mathcal{D}\backslash\mathcal{S}, t \in T, s \in S.\end{aligned} \tag{3.11}$$

Stockpile constraints are used to ensure only the amount of material within a stockpile can leave and to calculate the end-of-period stockpile inventories based on the incoming and outgoing material at a stockpile:

$$\sum_{j\in\mathcal{O}(i)} y_{i,j,t} \leq 1 \quad \forall i \in \mathcal{S}, t \in T. \tag{3.12}$$

$$\mathrm{STK}_{i,t,s} = \mathrm{STK}_{i,(t-1),s}\left(1 - \sum_{j\in\mathcal{O}(i)} y_{i,j,t}\right) + \sum_{g\in G} \kappa_{g,i,\tau}\left(\lambda_{g,i,t,s}^{\text{coarse}} + \lambda_{g,i,t,s}^{\text{fine}}\right)$$
$$\forall i \in \mathcal{S}, t \in T, s \in S. \tag{3.13}$$

$$\mathrm{MSTK}_{q,i,(t+1),s} = \mathrm{MSTK}_{q,i,t,s}\left(1 - \sum_{j\in\mathcal{O}(i)} y_{i,j,t}\right) + \sum_{g\in G} \kappa_{g,i,t}\left(\gamma_{q,g,i,t,s}^{\text{coarse}} + \gamma_{q,g,i,t,s}^{\text{fine}}\right)$$
$$\forall q \in Q, i \in \mathcal{S}, t \in T, s \in S. \tag{3.14}$$

Short-term operational constraints enforce soft constraints on production capacities and balance production at certain destinations by applying a lower bound. This includes mining, screening, and processing capacities:

$$\sum_{i\in\mathcal{D}} \sum_{g\in\mathcal{G}} \kappa_{g,i,t}(\lambda_{g,i,t,s}^{\text{coarse}} + \lambda_{g,i,t,s}^{\text{fine}}) - \delta_{i,t,s}^{+} \leq \mathrm{U}_{i,t} \quad \forall i \in \mathcal{M}, t \in T, s \in S \tag{3.15}$$

$$\sum_{i\in\mathcal{D}} \sum_{g\in\mathcal{G}} \kappa_{g,i,t}(\lambda_{g,i,t,s}^{\text{coarse}} + \lambda_{g,i,t,s}^{\text{fine}}) + \delta_{i,t,s}^{-} \geq \mathrm{L}_{i,t} \quad \forall i \in \mathcal{M}, t \in T, s \in S \tag{3.16}$$

$$\mathrm{STK}_{i,t,s} - \delta_{i,t,s}^{+} \leq \mathrm{U}_{i,t} \quad \forall i \in \mathcal{S}, t \in T, s \in S \tag{3.17}$$

$$\mathrm{STK}_{i,t,s} + \delta_{i,t,s}^{-} \geq \mathrm{L}_{i,t} \quad \forall i \in \mathcal{S}, t \in T, s \in S \tag{3.18}$$

$$\left(\mathrm{MASS}_{i,t,s}^{\text{coarse}} + \mathrm{MASS}_{i,t,s}^{\text{fine}}\right) + \delta_{i,t,s}^{-} \geq \mathrm{L}_{i,t} \quad \forall i \in \mathcal{D}\backslash\mathcal{S}, t \in T, s \in S \tag{3.19}$$

$$\left(\mathrm{MASS}_{i,t,s}^{\text{coarse}} + \mathrm{MASS}_{i,t,s}^{\text{fine}}\right) - \delta_{i,t,s}^{+} \leq \mathrm{U}_{i,t} \quad \forall i \in \mathcal{D}\backslash\mathcal{S}, t \in T, s \in S \tag{3.20}$$

$$\left(\mathrm{MASS}_{i,t,s}^{\text{coarse}} + \mathrm{MASS}_{i,t,s}^{\text{fine}}\right) + \delta_{i,t,s}^{-} \geq \mathrm{L}_{i,t} \quad \forall i \in \mathcal{D}\backslash\mathcal{S}, t \in T, s \in S. \tag{3.21}$$

Schedule compliance constraints mitigate the risk of the short-term production schedule forecasts deviating from long-term targets. The connection between short- and long-term scheduling is accomplished by introducing a set of soft constraints. The first set ensures that the total tonnage mined over short-term schedule meets long-term production targets as this is critical in mining to ensure future areas of extraction are available to be mined. Furthermore, a set of soft constraints are included to align the short-term forecasted production of property $q \in Q$ produced at processing facility $i \in \mathcal{P}$ with the targets from the long-term production schedule to ensure consistent metal production:

$$\sum_{t \in T} \sum_{i \in \mathcal{D}} \sum_{g \in \mathcal{G}} \kappa_{g,i,t} (\lambda_{g,i,t,s}^{\text{coarse}} + \lambda_{g,i,t,s}^{\text{fine}}) + \delta_s^{\text{TT},-} \geq \text{TT} \quad \forall s \in S \tag{3.22}$$

$$\sum_{t \in T} \left(\text{METAL}_{q,i,t,s}^{\text{fine}} + \text{METAL}_{q,i,ts}^{\text{coarse}} \right) + \delta_{q,i,s}^{\text{TP},-} \geq \text{TP}_{q,i} \qquad \forall q \in Q, i \in \mathcal{P}, s \in S. \tag{3.23}$$

Non-negativity constraints are included, and indices are omitted for clarity:

$$y, \delta, \lambda, \gamma \geq 0\,. \tag{3.24}$$

Lastly, mining reserve, block access and destination policy constraints for groups are included, as in Goodfellow and Dimitrakopoulos (2016). The following section discusses the solution method to the stochastic mathematical programming formulation presented.

3.4 Combining reinforcement learning with stochastic mathematical programming

3.4.1 Reinforcement learning background

In reinforcement learning, an agent interacts with an environment over a series of discrete timesteps (Sutton and Barto 2018). At each timestep the agent observes a state s_t and selects an action a_t using a learned policy π. The policy defines the mapping from a state to the probabilities of each possible action. After selecting an action, the environment responds by providing the agent with the next state s_{t+1} and a reward r_{t+1} that is used to guide the learner. The agent aims to maximize the total return by adapting the policy using a reinforcement learning algorithm. The return $G_t = \sum_{k=0}^{T} \gamma^k r_{t+k+1}$ from any given timestep is simply a function of the reward sequence where $\gamma \in (0,1]$ is the discount rate and T denotes the terminating time step. In this work, the timesteps represent production periods and the terminating timestep is the final production period.

There are several classes of reinforcement learning algorithms including value-based methods, policy gradient methods and actor-critic algorithms (Sutton and Barto 2018). These methods rely on the use of function approximators, such as neural networks, to generalize over states (Mnih et al. 2013; Mnih et al. 2015; Mnih et al. 2016; Haarnoja et al. 2018). Value based methods use reinforcement learning algorithms to approximate the value function or action-value function for each state and create a policy directly from the estimated values, possibly using a greedy or ϵ-greedy policy (Lillicrap et al. 2015). On the other hand, policy gradient methods directly parametrize the policy (Sutton et al. 1999). As a function of the state, the policy can be approximated with a neural network where the parameters are learned for which action to take given each state by maximizing the return (Mnih et al. 2016). In this work, the return relates to the value of the stochastic programming objective function and production scheduling decisions. Policy gradient methods are particularly useful as they can also learn continuous actions (Sutton and Barto 2018). This is advantageous over value-based method as it is difficult to define a policy using a value function for a large number of actions. Actor-critic algorithms combine aspects of both the previous methods. The actor learns a policy that maximizes cumulative return, and a critic learns to estimate the value function. The state-value function, defined as $V_\pi(\mathcal{s}) = \mathbb{E}_\pi[G_t | \mathcal{s}_t = \mathcal{s}]$, is the expected return from a given state $\mathcal{s}$ under policy π and is commonly used as the critic (Sutton and Barto 2018). Actor-critic methods combined with deep neural networks have proven to outperform experts in computer games and 3-D virtual environments (Mnih et al. 2016) and are applied in this work to learn a discrete-continuous policy for short-term production scheduling.

3.4.2 Modelling the agent-environment interface

A short-term production schedule is generated by optimizing the stochastic mathematical programming formulation introduced in Section 3. To optimize the model, an agent-environment interface is defined that allows a reinforcement learning agent to interact with a mining complex environment. A schematic of the agent-environment interface is shown in Figure 3.4. The agent is represented by a neural network policy and the environment is modeled using the stochastic mathematical programming formulation from Section 3. The environment provides feedback to the agent by providing it with states (contextual information related to the current production schedule) and rewards.

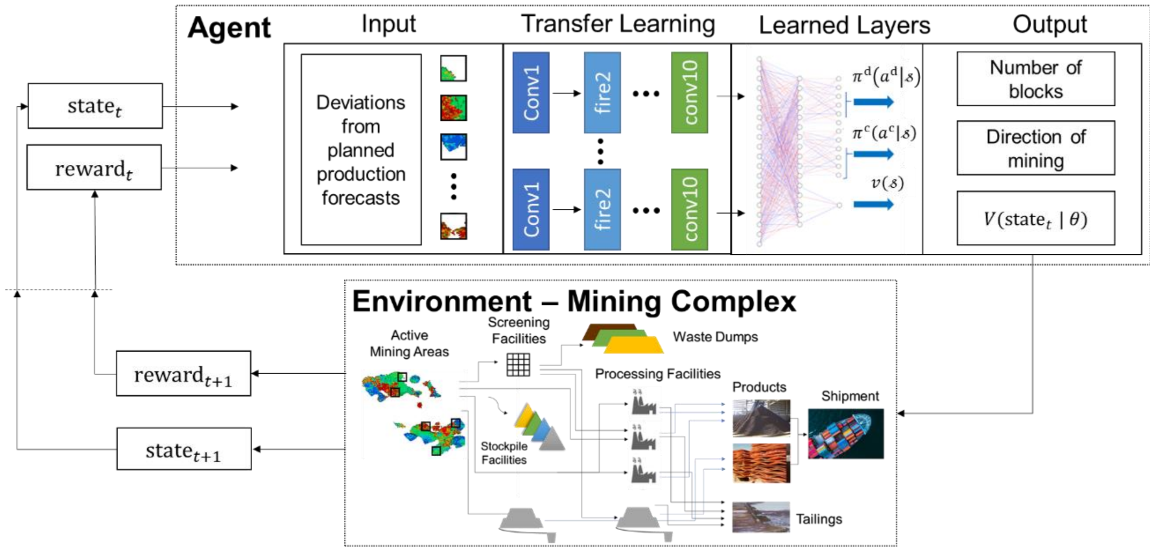


Figure 3.4. The proposed reinforcement learning agent-environment interface for mining complexes

An actor-critic reinforcement learning algorithm (Mnih et al. 2016) is applied to train the agent and directly learn a policy for determining the extraction sequence with a combination of discrete and continuous actions. The continuous actions determine the number of blocks to extract from each available access point and the discrete actions determine the direction of mining. A state-value function is also learned to predict the expected future return and used as the critic in the actor-critic reinforcement learning algorithm discussed in the subsequent section. The state-value function is denoted by $V(s_t; \theta_v)$ with parameters θ_v.

The reinforcement learning environment accounts for the n access points currently available in a mining complex. Access points are denoted AP_j for $j = 1, \dots, n$ and their corresponding geographical position is denoted POS_j. Mineable blocks nearest to each access point position POS_j are defined using a neighbourhood function $N(\text{POS}_\text{j})$. The material properties of those blocks within the neighbourhood provide the agent with contextual information that helps it decide how to mine at different access points to maximize the objective function of the stochastic programming formulation. These block properties make up a part of the mining complex state s that is feed to the reinforcement learning agent:

$$s = (\text{BP}, \text{DEV}, \text{CAP}) \tag{3.25}$$

where BP is the set of block properties, DEV is the set of deviations from long-term production targets and CAP represents the positive and negative deviations from short-term capacities.

An example of the components that make up the state are shown in Figure 3.5. The neighbourhood of blocks $N(\mathrm{POS}_j)$ at each access point $j = 1, \ldots, n$ are those within the outlined squares on the left of Figure 3.5.

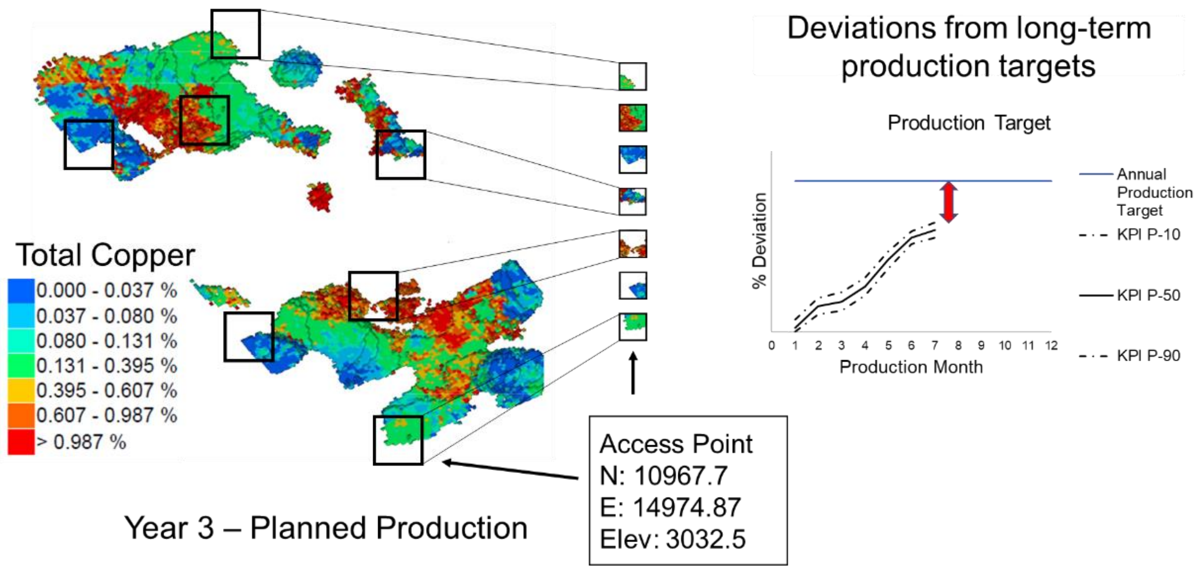


Figure 3.5. An example of the data collected from the mining environment during each timestep. On the left, the block properties (BP) considered at each access point within a neighbourhood (square) and on the right the deviations (DEV) from targets are shown. This state information is extracted directly from the parameters and recourse variables of the stochastic mathematical programming formulation.

A neural network policy is used to determine the blocks to be extracted at all access points in a mining complex. The policy $\pi(a|s;\theta)$, with parameters θ, maps the state of the mining complex to a discrete-continuous action that delimits the set of blocks $\mathcal{B}_t$ to be extracted at each access point in production period t. The discrete actions component a_j^{d} decides the mining direction (east, west, north, south, or radial) and the continuous actions a_j^{c} determine the number of blocks to mine at each access point $j = 1, \ldots, n$. An example of the combined discrete-continuous action for a single access point are shown in a simplified 2-D example in Figure 3.6. The blue blocks represent the neighbourhood of blocks to be mined by applying the discrete-continuous action at an access point during a single production period. The action space selects an extraction sequence that always satisfies the proposed model feasibility by ensuring that the slope constraints and reserve constraints are satisfied.

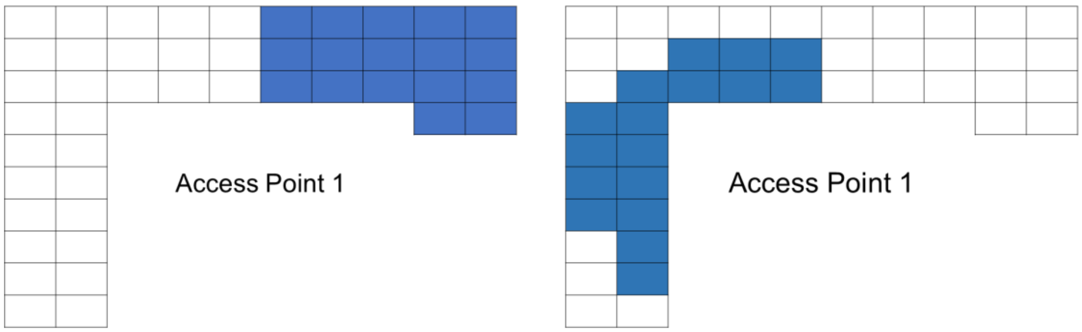


Figure 3.6. A simplified 2-D example of the parametrized policy using a continuous and discrete action space shows an east action selection (left) versus a west action selection (right). The continuous variable allows for seventeen blocks to be extracted in blue at each access point.

Given the extraction sequence determined by the reinforcement learning agent and initial stockpile inventories, the downstream destination policy, stockpile and preconcentration decisions can now be optimized by solving the stochastic mathematical programming formulation in Section 3. The objective function expressed in Eq. 3.1 is used to evaluate the agents action subject to two sets of additional constraints to the original formulation that fix the extraction sequence decisions that are determined using the reinforcement learning agent:

$$x_{b,\tau} = 1 \;\; b \in \mathcal{B}_\tau, \tau = 1, \dots, t. \tag{3.26}$$

$$x_{b,\tau} = 0 : \;\; b \notin \mathcal{B}_\tau, \tau = 1, \dots, T. \tag{3.27}$$

Eq. 3.26 and 3.27 show the connection mechanism between the reinforcement learning agent and the stochastic mathematical programming formulation. The optimal objective value, denoted by objective$_t^*$ is obtained by maximizing Eq. 3.1 subject to constraints 3.26 and 3.27 using CPLEX v20.1. The reward r_{t+1} generated by the mining complex environment informs the agent of its performance during each timestep and is simply the difference between the current optimal objective and the one obtained on the previous timestep:

$$r_{t+1} = \text{objective}^*_{t+1} - \text{objective}^*_t. \tag{3.28}$$

Defined in this way, the reward is the incremental improvement of the short-term production schedule for scheduling the most recent production period. The cumulative return obtained for solving each production schedule equates to the optimized value of the objective function in the stochastic mathematical programming formulation given the extraction sequence selected by the agent.

The mining complex environment also provides new states to the agent after each action is taken. Each component of the state is updated with information that comes directly from the stochastic programming formulation in Section 3, either from the parameters of the problem, related to the simulated block model, or the optimized decision variables.

After each action is taken by the agent, the environment responds by first updating the POS_j for each access point $j = 1, \dots, n$ to the nearest unmined block, which modifies the block properties closest to access point j:

$$\text{BP} = \left\{\left(g_{b,q,s}, \text{RF}_{b,q,i,s}, m_{b,s}\right): b \in N\left(\text{POS}_j\right), \forall j = 1, \dots, n, q \in Q, i \in, s \in S\right\}. \tag{3.29}$$

Deviation (DEV) and capacity (CAP) components are updated with the deviation variables retrieved from the optimized stochastic mathematical programming formulation:

$$\text{DEV} = \left\{\left(\delta_s^{\text{TT},-}, \delta_{i,s}^{\text{TP},-}\right): i \in \mathcal{P}, s \in S\right\}. \tag{3.30}$$

$$\text{CAP} = \left\{(\delta^-_{i,t,s}, \delta^+_{i,t,s}): i \in \{\mathcal{M} \cup \mathcal{D}\}, t \in T, s \in S\right\}. \tag{3.31}$$

The agent-environment interface explained here is used to learn a policy for optimizing the short-term production schedule. In the following section, an explanation of the steps taken to improve the policy are introduced.

3.4.3 Optimizing a discrete-continuous policy with actor-critic reinforcement learning

The states, actions and rewards generated by interacting with the mining complex environment are used for optimizing the policy parameters θ of a neural network for selecting extraction sequence decisions, described in Appendix A. In this work, an advantage actor-critic (A2C) reinforcement learning approach (Mnih et al. 2016) is applied to directly learn a parameterized policy for a set of discrete and continuous actions $\pi^{\mathrm{d}}(a^{\mathrm{d}}|\mathcal{s};\theta)$ and $\pi^{\mathrm{c}}(a^{\mathrm{c}}|\mathcal{s};\theta)$, respectively. The A2C algorithm requires maintaining an estimate of the value function $V_{\pi}(\mathcal{s};\theta)$ under policy π. To optimize the discrete-continuous policy, independence is assumed between each action (Delalleau et al. 2019; Neunert et al. 2020) such that:

$$\pi(a|\mathcal{s};\theta) = \prod_{j=1}^{n} \pi_j^{\mathrm{d}}\left(a_j^{\mathrm{d}}\middle|\mathcal{s};\theta\right) \prod_{j=1}^{n} \pi_j^{\mathrm{c}}\left(a_j^{\mathrm{c}}\middle|\mathcal{s};\theta\right) \tag{3.32}$$

where a_j^{c} and a_j^{d} are the continuous and discrete actions for each access point $j = 1, \dots, n$. The components of the continuous policy $\pi_j^{\mathrm{c}}(a|\mathcal{s};\theta)$ control the quantity of blocks mined at each access point and are represented as a normal distribution for $j = 1, \dots, n$:

$$\pi_j^{\mathrm{c}}\left(a_j^{\mathrm{c}}\middle|\mathcal{s};\theta\right) = \mathcal{N}\left(\mu_j(\mathcal{s};\theta), \sigma^2\right) \tag{3.33}$$

where the standard deviation σ is a parameter and μ_j is the output of a neural network. Furthermore, the discrete policy $\pi_j^{\mathrm{d}}\left(a_j^{\mathrm{d}}\middle|\mathcal{s};\theta\right)$ is represented with a softmax activation function on the neural network output, providing a categorical distribution over each possible direction.

The discrete-continuous action space is used to simultaneously optimize the extraction sequence at each access point with an approach similar to those decisions being made at a mining operation. The operations planning teams must determine the direction to mine given the current available access points. Then, a balancing of the required material properties is determined by varying the number of materials taken from each area. The combination of the discrete-continuous actions automates this decision-making process, while also enforcing minable extractions sequences.

The advantage actor-critic reinforcement learning algorithm applied is an on-policy method that uses several parallel instances of the mining complex environment E and is a variant of the asynchronous approach discussed in Mnih et al. (2016). The agent interacts with parallel environments to generate different experiences depending on the actions selected. This allows for the data to be generated from several timesteps and helps decorrelate the updates of the neural network. The policy parameters θ are incremented in the direction of the approximate policy gradient [19], which is adapted here to consider a discrete-continuous policy:

$$\Delta\theta = \nabla_{\theta}\left(\psi_c \log\left(\pi^{\mathrm{c}}(a_t^{\mathrm{c}}|s_t;\theta)\right) + \psi_d \log\left(\pi^{\mathrm{d}}\left(a_t^{\mathrm{d}}\middle|s_t;\theta\right)\right)\right)\left(G_t - V(s_t;\theta)\right) + \psi_e \nabla_{\theta}\mathcal{H}\left(\pi_d\left(a_t^{\mathrm{d}}\middle|s_t;\theta\right)\right). \tag{3.34}$$

To prevent early convergence to a single discrete action, an entropy term $\mathcal{H}(\cdot)$ is added to adequately explore the discrete action component. Parameter ψ_e controls the contribution of the entropy regularization term, while parameters ψ_c and ψ_d emphasize the importance of each component in the discrete-continuous action space. The value function is optimized to minimize the prediction error with approximate gradient descent:

$$\Delta_{\theta_v} = \nabla_{\theta_v}\left(G_t - V_{\pi}(s_t;\theta_v)\right)^2. \tag{3.35}$$

Gradients are accumulated over the parallel environments and used to update the neural network weights and optimize the policy. Pseudocode for the actor-critic reinforcement learning approach is shown in Algorithm 1.

Algorithm 1 Actor-critic reinforcement learning for short-term production scheduling

1: **procedure** A2C($E, T_{\max}$)
2: // E is a set of mining complex environments
3: // Assume environment specific counters t are initialized to 1
4: // $T_{\max}$ is the number of updates
5: Initialize θ, θ_v ▷ Initialize policy and value function parameters
6: Initialize $d\theta, d\theta_v$ ▷ Initialize policy and value function gradients
7: $T_g = 0$ ▷ Initialize global shared counter
8: **repeat**
9: $d\theta \leftarrow 0$ and $d\theta_v \leftarrow 0$ ▷ Reset gradients
10: **for each** $e \in E$ **do**
11: $t_{\text{start}} = t$ ▷ Update to counter value for environment e
12: Get state s_t ▷ Get current state from environment e
13: **repeat**
14: Sample action a according to policy $\pi(\cdot|s_t;\theta)$
15: Update extraction decisions with action a
16: Solve stochastic program (Eq 3.1-3.24) with branch-and-cut
17: Compute reward r using Eq. 3.28
18: Update new state s_{t+1} with Eq. 3.29-3.31
19: $t \leftarrow t+1$
20: $T_g \leftarrow T_g + 1$
21: **until** terminal s_t or $t - t_{\text{start}} == t_{\max}$
22: $G = \begin{cases} 0, & \text{for terminal } s_t \\ V(s_t;\theta_v), & \text{otherwise} \end{cases}$
23:
24: **for each** $i \in \{t-1, \ldots, t_{\text{start}}\}$ **do**
25: $G \leftarrow r_i + \gamma G$
26: //Accumulate gradients using Eq. 3.34, 3.35
27: $d\theta \leftarrow d\theta + \nabla_\theta \log \pi(a_i|s_i;\theta)(G - V(s_i;\theta_v)) + \nabla_\theta \mathscr{H}(\pi^d(a_i^d|s_i))$
28: $d\theta_v \leftarrow d\theta_v + \nabla_{\theta_v}(G - V(s_i;\theta_v))^2$
29: **end for**
30: **end for**
31: $\theta \leftarrow \theta + d\theta$ ▷ Update parameters θ
32: $\theta_v \leftarrow \theta_v + d\theta_v$ ▷ Update parameters θ_v
33: **until** $T_g > T_{\max}$

Algorithm 1 is summarized as follows. A policy and value function are initialized with random and pretrained parameters using the neural network architecture described in Appendix A (line 5). Then, the gradients and global counter are initialized (lines 6-7). While the global counter is less then the maximum number of iterations ($T_{\max}$), the reinforcement learning agent interacts with the mining complex environment by generating new experiences (states, actions, and rewards) and accumulating gradients over several mining complex environments. The reinforcement learning agent samples a discrete-continuous action in line 14 from the learned policy. The action represents the direction and quantity of blocks to be mined. A set of blocks are selected for extraction given the direction and quantity of material to be mined while obeying slope constraints and reserve constraints. The blocks selected are used to update the constraints in Eq. 3.26 and 3.27 (line 15). Then, the stochastic mathematical programming formulation (Eq. 3.1-3.24) is optimized with a branch-and-cut strategy (implemented in CPLEX) given the updated extraction sequence decisions (line 16). Given the optimal decisions for the remaining decisions variables and the optimal objective function value, the reward can be computed with Eq. 3.28 (line 17) and used to update the policy and value function parameters. In line 18, the state of the reinforcement learning agent is updated by moving the agent the agent to the nearest available block for extraction updating the position of the agent and reconstructing the state by collecting the nearby information with Eq. 3.29-3.31. Lines 21-27 follow the policy and value function update for a typical actor critic reinforcement learning agent as discussed in (Mnih et al. 2016; Sutton and Barto 2018). Lastly, lines 30-31 update the neural network parameters. The proposed reinforcement learning algorithm and corresponding stochastic programming formulation are tested in the following section at a large-scale mining complex.

3.5 Case study at a copper mining complex

The proposed reinforcement learning framework is applied at a large-scale copper mining complex to optimize a monthly short-term production schedule for a single production year (12 months).

3.5.1 Overview of the mining complex

The mining complex, presented in Figure 3.1, is comprised of two open-pit mines, a sulphide stockpile, a screening preconcentration facility, three processing plants (mills), a sulphide bioleach, an oxide leach and a waste dump destination. The two open-pit mines supply materials to the processing destinations within the mining complex to produce copper concentrate and copper cathode products. There are two main material types: sulphides and oxides. Sulphide materials extracted from the mines can be sent to the screening preconcentration facility, sulphide stockpile, sulphide bioleach, processing plants or waste dump destination. Whereas oxide materials can be sent to the oxide leach, oxide stockpile or the waste dump. The screening preconcentration facility separates materials with beneficial grade-by-size response and can send the coarse and fine size fractions to either the waste dump, sulphide bioleach or processing plant destinations. The uncertainty of the mineral deposits are quantified using a set of 15 stochastic simulations of the material properties including soluble copper, total copper and the material response rank (Albor Consuegra and Dimitrakopoulos 2009). Soluble copper represents the fraction of total copper recoverable by leaching destinations. The grade-by-size considerations for preconcentration are simulated using a dataset of drilled core samples from the mineral deposit that are used to spatially simulate the areas of the deposit that have varying response rank using sequential indicator simulation (Remy et al. 2009).

The proposed reinforcement learning algorithm is used to determine a monthly short-term production schedule for twelve production months that maximizes profit while minimizing short- and long- term technical risk. There are five active access point locations with ramp and bench access in the two mines, each of them can supply material to the downstream destinations. All the components in the mining complex have a production capacity and operating cost ($/t). The cost related parameters for each location in the mining complex are shown in Table 3.1. In addition, the copper selling price is $2.24/lb. The parameters are scaled for confidentiality reasons. Table 3.2 shows the capacities of each location along with their corresponding penalty cost for deviations and applied metal recovery. The penalty costs are first estimated based on the unit cost of deviating from a capacity or forecast of interest, however, the parameters are modified experimentally by running the optimization model multiple times to retrieve a desired risk profile for necessary constraints (Benndorf and Dimitrakopoulos 2013). The desired risk profile may differ depending on a mining enterprises risk tolerance. Furthermore, the forecasted long-term production targets are considered when scheduling the 12-month short-term schedule. In this case, 525.2 million tonnes of material are extracted in a way that ensures accessibility to mining areas planned in future years. It also ensures that the forecasted operating profit of $2.9B is obtained over the 12 production months.

Table 3.1. Relative economic parameters for mining, processing, screening, rehandling costs and metal price

Location (i)	Economic Parameters ($c_{i,t}$)
Mining ($/t)	1.00
Mill 1 ($/t)	7.30
Mill 2 ($/t)	7.28
Mill 3 ($/t)	7.25
Oxide leach ($/t)	6.64
Bioleach ($/t)	2.10
Screening cost ($/t)	0.23

Table 3.2. Operational capacities, penalty costs and metal recovery

Location (i)	Capacity x 10^7 tonnes ($U_{i,t}$)	Penalty Cost ($pc^{+}_{i,t}$)	Recovery % ($r_{\text{copper},i,t}$)
Mine	54.6	500	-
Mill 1	3.4	200	83.0
Mill 2	2.3	175	80.4
Mill 3	1.3	175	82.6
Oxide Leach	1.5	175	65.0
Bioleach	5.0	10	27.0
Screening Facility	0.4	200	-
Sulphide Stockpile	10.0	10	-
Oxide Stockpile	10.0	10	-

3.5.2 Reinforcement learning parameters

The maximum number of blocks that can be extracted per period from each access point in the mining complex is 1731 blocks, based on the mining capacity. The stockpile is initialized with the quantity and quality of material available prior to the first planned production period. For the advantage actor-critic reinforcement learning approach a step size of 10^{-4} is used with the Adam optimizer (Kingma and Ba 2014). This is based on testing with learning rates between 10^{-2} and 10^{-5}. A slower learning rate improved the overall performance and maintained a stable policy. The parameters ψ_c and ψ_d are set to 1 and 0.25, respectively to manage the magnitudes of the discrete and continuous probabilities in the policy update. Higher values of ψ_e forced the probabilities of each discrete action to stay too close to each other, leading to a very exploratory policy. Therefore, ψ_e is set to 0.01 to encourage exploration, however, the policy will still adapt to those actions that maximize the expected return of the reinforcement learning agent. A limitation of applying reinforcement learning algorithms is determining the appropriate parameters for learning, which often requires a grid search and hyper-tuning. The parameters provided performed best for the given case study, however, different parameters may perform better on other case studies.

3.5.3 Base case production schedule

In this section, a base case production schedule is presented that defines a schedule for the mining complex without considering preconcentration. The extraction sequence, destination policy and stockpiling decisions are optimized together without screening facilities. Then, Section 5.4 considers the additional benefit of considering preconcentration via screening in the optimization formulation. The results showcase the incremental improvement of simultaneously optimizing additional decisions in a mining complex. In the risk profiles to follow, the green hashed lines represent the forecasted production based on the long-term production schedule, while the orange lines represent the capacities of the different destinations. The P-10, P-50, and P-90 represent the percentiles, which summarize the impact of material uncertainty on the production forecasts (i.e., 10% of forecasts fall below the P-10). Both of the example cases are run with an Intel® i7-8700 CPU and 32GB of memory.

The results obtained from optimizing the base case mine production schedule are shown in Figure 3.7, 3.8 and 3.9. The production forecasts show that there is a large risk of deviating from production forecasts and capacities largely due to the limitations of accessing material through available access points. In Figure 3.7, the amount of material placed at the oxide leach facility shows there are deviations from forecasts in month 4, 5, 7 and 9, while earlier years do not receive as much material as needed. This is largely due to the amount of oxide material being mined at the surface, which can not be processed using alternate methods. The sulphide bioleach has similar issues after month 6, however, the additional available capacity at this location can be used to manage the risk of deviating from targets as there is unused capacity in the long-term plan. Deviations also occur at the mill process plants shown in Figure 3.8. Mill 1 and Mill 2 are satisfied quite well over most production months. However, Mill 3 has large deviations from the operational capacity in period 1, 4 and 11. This is due to underfeeding the mills in some production months and only getting access to mill quality material in later production periods. These deviations can not be mitigated as the stockpile has also reached its capacity (Figure 3.9). The total cumulative operating cashflows exceed $3B. This satisfies long-term requirements; however, the schedule is optimistic due to the additional feed at the oxide leach and Mill 3 that likely can not be processed.

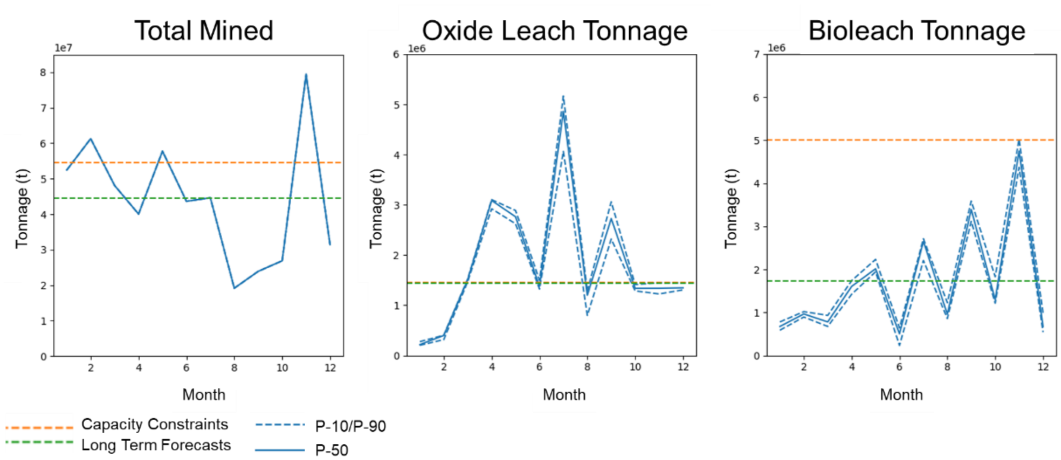


Figure 3.7. Base case production forecast and risk analysis for total material mined and leaching facilities.

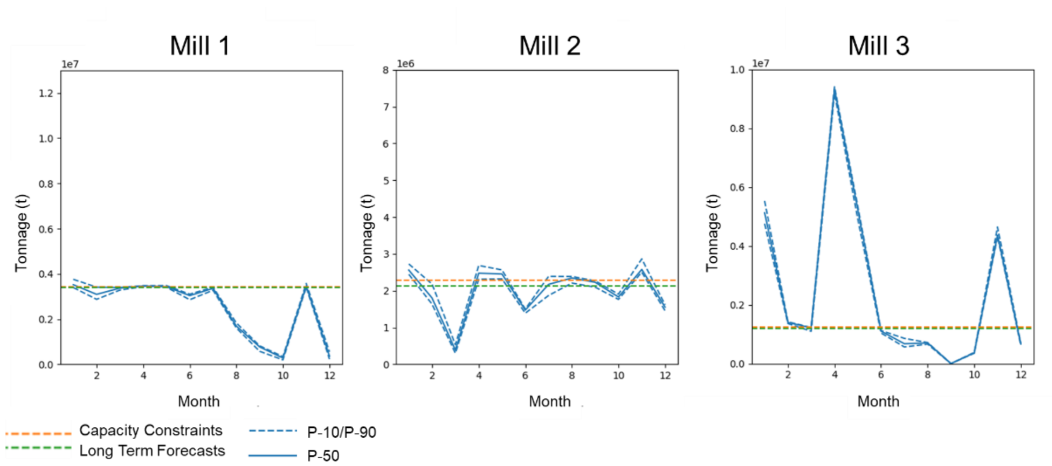


Figure 3.8. Base case production forecast risk analysis for the three mill process facilities.

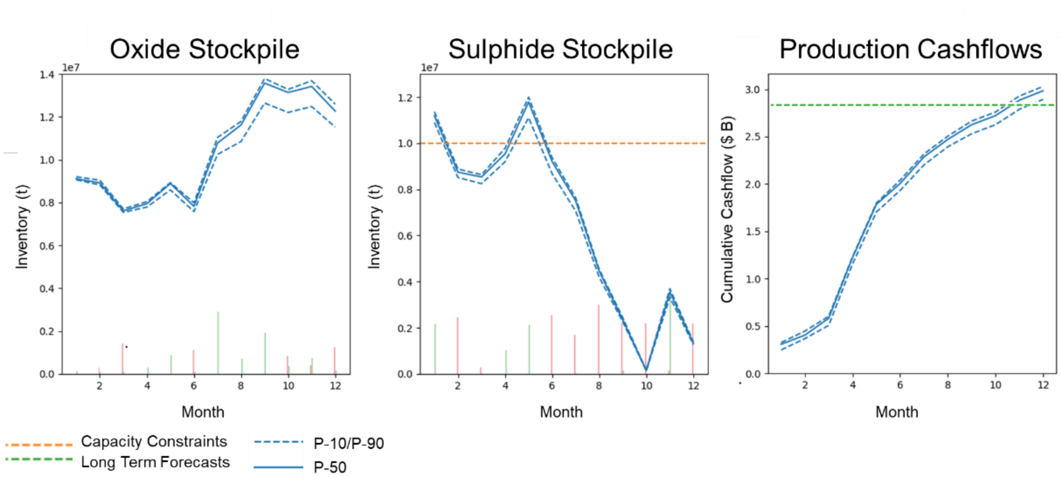


Figure 3.9. Base case production forecast and risk analysis of stockpile levels and production cashflows.

3.5.4 Production schedule with integrated preconcentration

In this section, the results from simultaneously optimizing the mining complex with preconcentration are shown. Figure 3.10 shows that the extraction sequence changes drastically, as the total amount of material mined changes for each production month. This is caused by the schedule adapting to balance the amount of material mined that is suitable for screening. The additional operational selectivity provided by screening materials based on its grade-by-size behaviour reduces the magnitude of the deviations at several components when comparing it to the base case production schedule. It reduces deviations from the oxide leach capacity in Figure 3.10 and violations from capacities at the mills shown Figure 3.10 and Figure 3.11, respectively. In addition, the stockpiled capacity is only violated in the first production month (Figure 3.12). Additional stockpiling capacity or a mill expansion will need to be investigated to manage the additional production expected at Mill 3 that will likely not be processed this year and instead will require rehandling in future years. Further value is obtained by simultaneously optimizing the schedule and integrating preconcentration via screening leading to a $140 M increase over the base case schedule. The improved short-term schedule generates higher cashflows while improving the ability to manage short-term capacities and still meet long-term targets. Lastly, a notable result is the substantial use of the screening facility during the first eight months of production, shown in Figure 3.12, as it helps balance production and allows for the operation to capitalize on synergies that were not apparent in the long-term schedule.

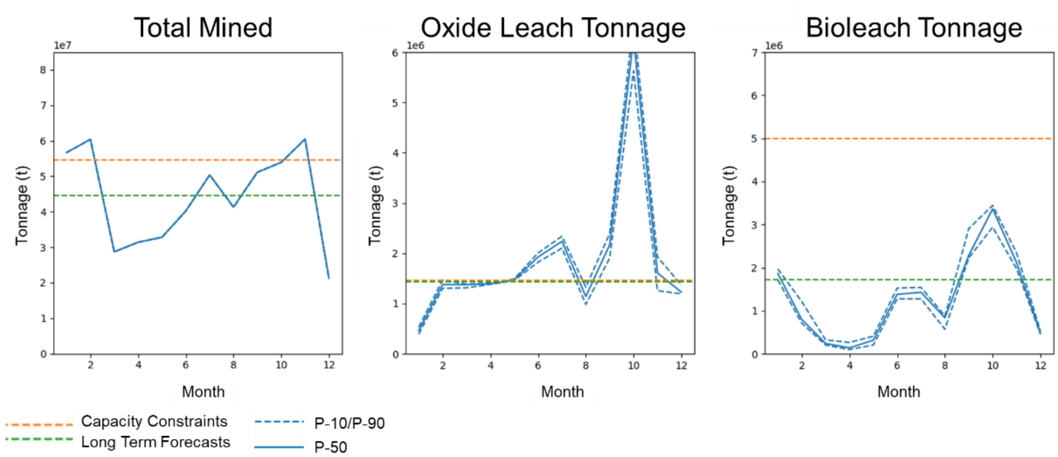


Figure 3.10. Production forecasts with preconcentration and risk analysis for total material mined and leaching facilities.

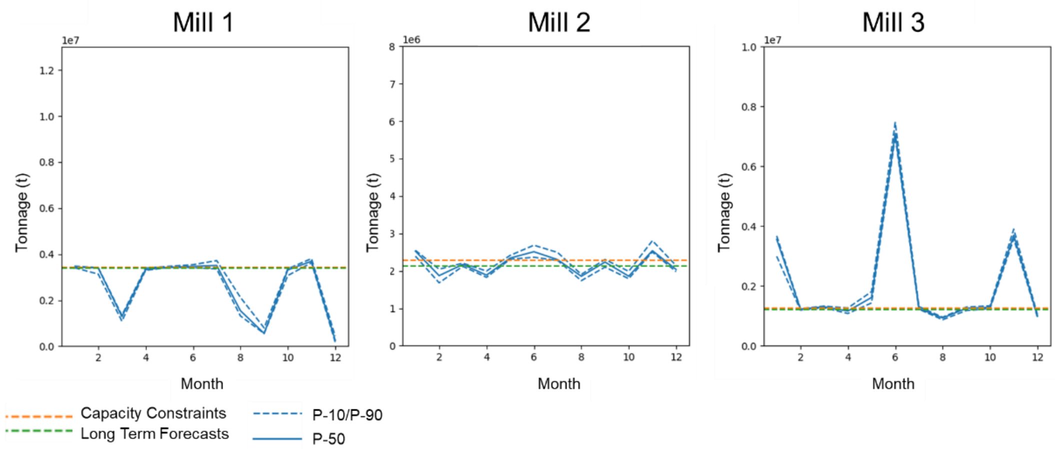


Figure 3.11. Production forecast with preconcentration and risk analysis for the three mill processing facilities.

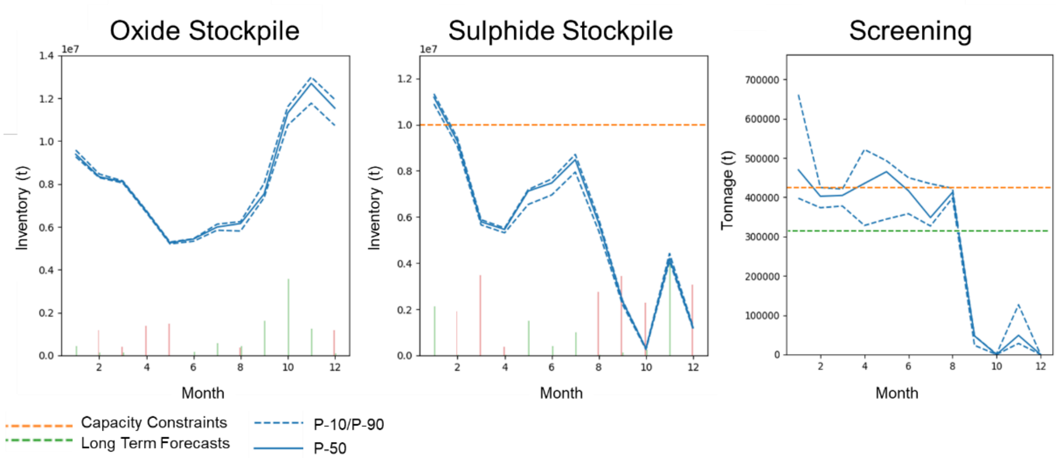


Figure 3.12. Production forecast with preconcentration and risk analysis of stockpile levels and screening facility production.

3.5.5 Discussion

In short-term production scheduling, the heterogeneity of materials and the ability to access material of high quality is a major factor in meeting the long-term forecast. Operational access points are an important aspect as they simulate the behaviour of operating mines and their ability to access material. Material accessibility is one of the key reasons why long-term production targets are difficult to obtain. When access points are considered in the optimization process, forecasts are closer to reality and scheduling problems can be addressed by considered stockpiling strategies or by adjusting the production plan. Introducing preconcentration into the simultaneous stochastic optimization framework helped alleviate some of the deviations from capacity, while adding additional value. It is not a perfect solution and further considerations should be investigated to find additional approaches for optimizing the short-term schedule. The differences in the base case schedule compared to the schedule that considers preconcentration are shown in Figure 3.13. It is clear that the extraction sequence changes when different scheduling considerations are included in the stochastic programing formulation.

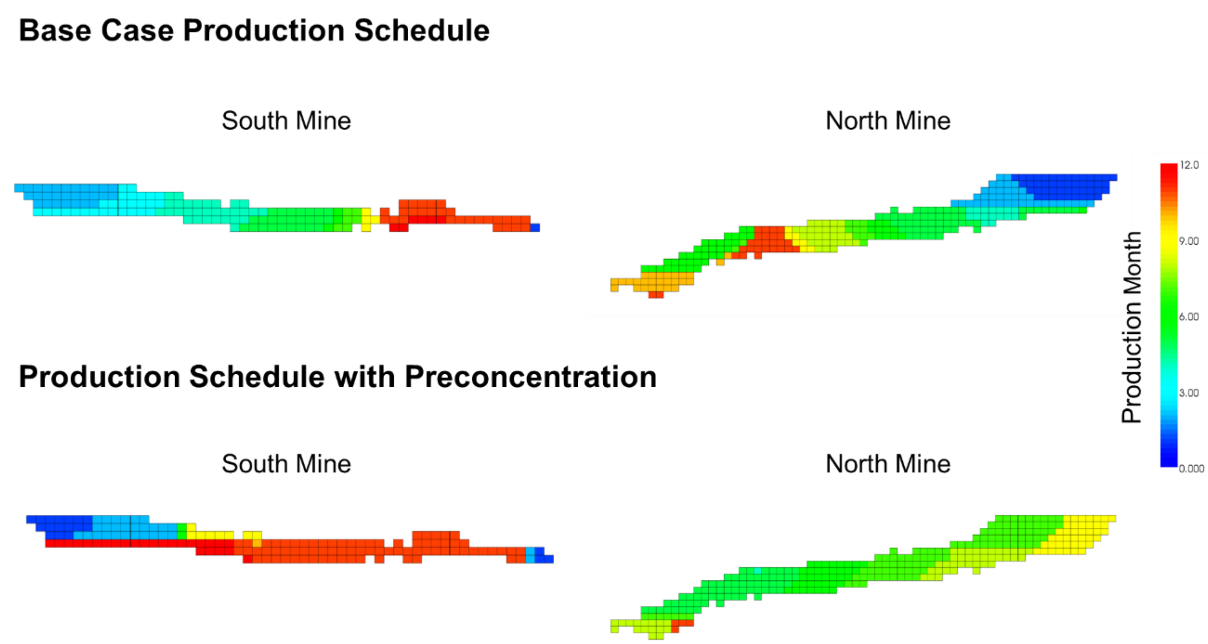


Figure 3.13. N-S cross section of the monthly extraction sequence for the base case extraction production schedule and production schedule that with preconcentration.

Another important aspect of this case study is the discrete-continuous action space used for learning to make extraction sequence decisions. Parametrizing a policy using neural networks and applying reinforcement learning to learn a policy for production scheduling demonstrated some beneficial results. First of all, the extraction sequence decisions directly considered operational access point locations in the mining complex, which led to operational extraction sequences. The extraction sequence for the results considering preconcentration can be seen clearly in Figure 3.14. The blue areas show the advancement in the first production month given the provided access points. The areas are continuous and mineable making it reasonable to develop ramps and benches, a critical aspect for operations. Reinforcement learning is an efficient method for ensuring minable shapes and eliminates the need for additional constraints (Matamoros and Dimitrakopoulos 2016; Quigley and Dimitrakopoulos 2020).

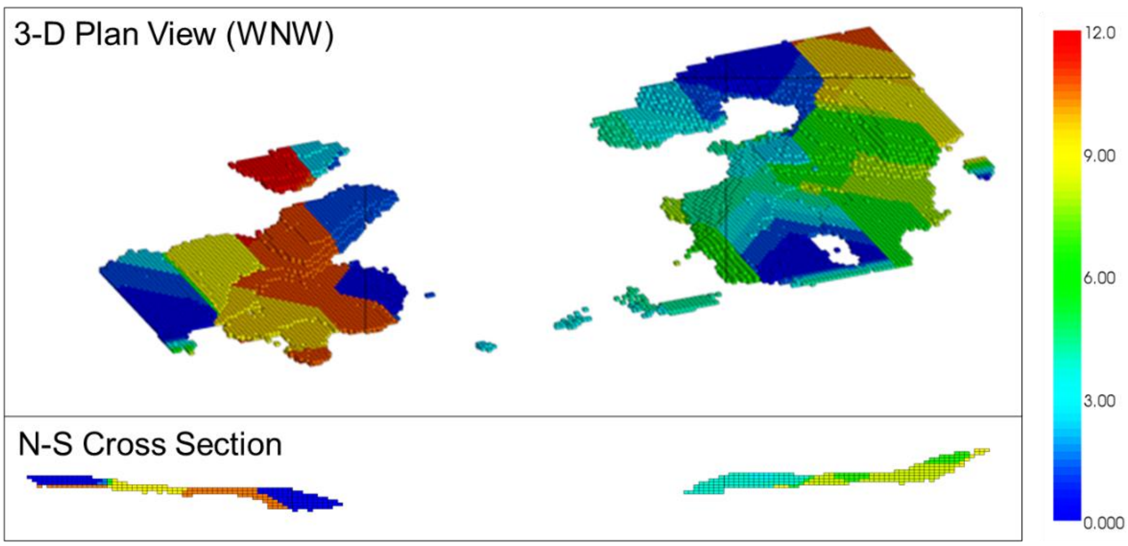


Figure 3.14. The adaptive production schedule with preconcentration; a 3-D Plan view of the extraction sequence looking WNW and a N-S Cross Section.

Finally, the performance of the proposed advantage actor-critic algorithm is shown in Figure 3.15. This is the tested performance of the reinforcement learning after each update to the policy over 10 runs of the A2C algorithm. The agent constantly explores new actions to develop the best strategies for optimizing a production schedule. After 500 policy updates a stable policy is achieved, the variance in the agents return decreases and returns are consistent. The training time required to achieve a stable policy requires 6.7 ± 0.25 h, however, once the reinforcement learning agent is trained it takes less than 1 minute to retrieve an optimized schedule using the learned neural network parameters. An interesting aspect of reinforcement learning is that the policy parameters are learned. These parameters can be saved and potentially reused on future optimization instances for the same mining complex. The agent in this case learns from information surrounding the access points and is not mastering the exact problem with full information but, learning on how to behave based on its progression from the access points. Designed in this way, the policy parameters can potentially be reused in the same mining complex in later months to reoptimize the schedule in light of new information. Learning the policy parameters could lead to benefits for optimization approaches that combine reinforcement learning with stochastic mathematical programming.

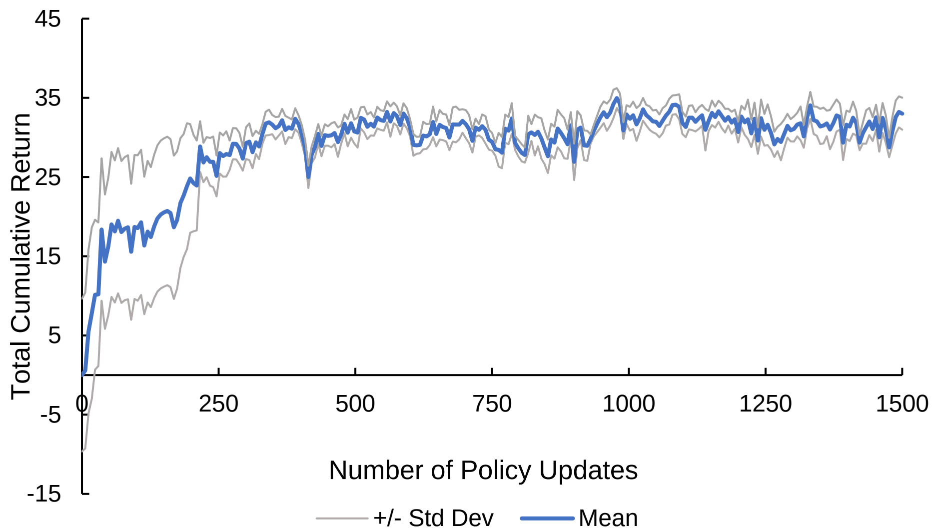


Figure 3.15. Actor critic reinforcement learning agent performance for short-term production scheduling.

3.6 Conclusions

A stochastic mathematical programming formulation for short-term production scheduling is presented with integrated preconcentration. Material grade-by-size behaviour and a coarse liberation model are used to model the separation of material at screening preconcentration facilities. The integration of preconcentration provides opportunities to capture additional value by separating coarse and fine material with a beneficial response rank prior to crushing and grinding in a process plant. To solve the new stochastic mathematical programming formulation, a novel approach is developed that combines reinforcement learning with stochastic programming. As a result, a reinforcement learning agent learns to optimize the short-term production schedule. A stable policy is learned for determining operational extraction sequences using a discrete-continuous action space for action selection, which ensures operational ramp and bench access. The framework is tested at a copper mining complex where it shows the direct benefits of simultaneously optimizing the production schedule with preconcentration decisions. This demonstrates a $140 M increase in annual cashflows, while helping manage short and long-term risk.

Future work should aim to improve the action space in the reinforcement learning approach to have a more diverse set of directions to mine and consider underground scheduling requirements. Furthermore, additional preconcentration methods or a combination of methods applied at different stages in the mining complex should be considered to increase the value of a mining complex and eliminate the unnecessary need to process excess waste. This includes opportunities to consider additional preconcentration techniques, for instance, bulk sorting and particle sorting. Lastly, future work should investigate the reuse of the learned policy for new areas in the same mining complex to see if this tool can be used for rapidly optimizing the short-term schedule and assessing its performance with respect to the annual production forecasts.

3.7 Appendix A: Neural network

An architectural choice is made that utilizes transfer-learning, an area common in image classification (Pan and Yang 2009). Transfer learning uses knowledge obtained from performing one task such as classifying an image and applies those learnings to perform a related task. The learned neural network parameters θ of the reinforcement learning agent can be reduced significantly by using a pre-trained neural network to extract high-level features from an image. In the reinforcement learning framework proposed, an image is constructed with the spatial arrangement of block properties nearest to each access point with the simulated values of the block grades as input. The inputs are transformed into images with dimensions of 224 by 224 pixels and input into a pre-trained convolutional neural network with 18 layers (Iandola et al. 2016). These layers are not trained, but solely used for feature extraction. Two additional fully-connected layers are added to the pretrained network with leaky ReLU activation (Maas et al. 2013). Each layer has 256 neurons. The final outputs of the neural network are then decoupled into separate network components, one for each access point AP_j, $j = 1, \ldots, n$, and their corresponding continuous and discrete policies. The continuous output components use tanh activation and the discrete policy uses softmax activation. Lastly, a single neuron is used to approximate the value function $V_\pi(s)$ with ReLU activation. The corresponding network for a single access point is shown in Figure 3.16 and can be expanded to select the extraction sequence for several access points. The neural network parameters are trained with reinforcement learning using Algorithm 1.

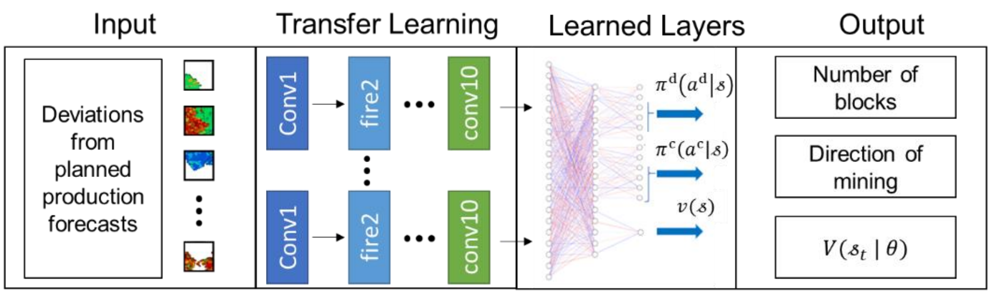


Figure 3.16. Discrete-continuous reinforcement learning neural network.

3.8 Chapter discussion and next steps

Chapter 3 presents a stochastic programming and reinforcement learning framework for optimizing the short-term production schedule in mining complexes. The short-term optimization framework integrates preconcentration decisions. Grade-by-size material behaviour is quantified with a response rank parameter and stochastic orebody simulations. Screening facility activities are modeled by allowing for the separation of coarse and fine materials at the preconcentration destinations. Furthermore, an innovative solution approach that applies an actor-critic reinforcement learning approach with a discrete continuous action space is developed to optimize the stochastic programming formulation. The reinforcement learning agent generates practical mining sequences based on the action-space defined and learns a policy for optimizing the short-term production schedule considered multiple active mining areas. The advantage of this method is that the learned policy can be applied to find different short-term production schedules in the same mineral deposit and provides a useful heuristic for evaluating short-term production schedules. In Chapter 4, the proposed optimization approach is integrated into a simultaneous stochastic optimization framework for jointly optimizing short-term production schedules. The development of this tool provides the initial building block for connecting planning horizons in industrial mining complexes.

4 Connecting planning horizons in mining complexes with reinforcement learning and stochastic programming

Connecting short- and long-term production schedules in mining complexes is essential to ensure that a long-term production schedule is achievable at shorter timescales. Previous research that addresses optimizing mining complexes under uncertainty primarily focuses on simultaneously optimizing different components in the mining complex to capitalize on advantageous synergies. Typically, short- and long-term production schedules are optimized separately in a number of stages. This poses a risk of schedule misalignment, which can adversely affect the economic outcome of a mining complex and the ability to meet long-term production forecasts at shorter timescales. Therefore, an integrated approach that considers short- and long-term production scheduling is proposed that leverages stochastic mathematical programming and reinforcement learning to connect planning timescales. The solution approach is tested in large operating copper mining complex and demonstrates significant improvements in the resulting production and financial forecasts.

4.1 Introduction

A mining complex is an integrated supply chain comprised of several components designed to extract, transport and transform raw materials into a set of valuable products for delivery to customers and the market (Barbaro and Ramani 1986; Hoerger et al. 1999a; Hoerger et al. 1999b; Whittle 2007; Pimentel et al. 2010). Each mining complex may contain several interconnected mines, stockpiles, crushers, leach pads, process plants, preconcentration and waste facilities (Hustrulid et al. 2013; Bowman and Bearman 2014). An example of a mining complex is shown in Figure 4.1. The primary source of technical risk in mining complexes relates to the quantity and quality of the material within the mineral deposits to be mined. The uncertainty and variability of the unknown material properties largely impacts production scheduling decisions. Therefore, it is vital to quantify the uncertain sources of material, represented using stochastic orebody simulations (Goovaerts 1997; Rossi and Deutsch 2014), and to manage risk within a production scheduling optimization framework.

Several stochastic mathematical programming formulations have been developed for separately optimizing short- and long-term production schedules in mining complexes (Montiel and Dimitrakopoulos 2015; 2018; Goodfellow and Dimitrakopoulos 2016; 2017; Both and Dimitrakopoulos 2020; 2023b; Dimitrakopoulos and Lamghari 2022). These formulations manage risk by accounting for material uncertainty and overcome the limitations of traditional optimization formulations that use estimated (average type) models to represent the material supply. Additionally, the joint behaviour of mining complex components are modeled, allowing for the simultaneous optimization of the extraction sequence, destination policy, processing stream, equipment allocation, capital investment and other major production scheduling decisions (Saliba and Dimitrakopoulos 2019; 2020; Levinson and Dimitrakopoulos 2020b; Both and Dimitrakopoulos 2021). Managing supply uncertainty and jointly optimizing components in a mining complex can significantly increase project value; however, the optimization process is particularly challenging due to the large and non-linear formulation required to model a mining complex with uncertainty. For this reason, metaheuristics are required to simultaneously optimize production scheduling decisions for large-scale mining complexes, which are intractable to solve using conventional mathematical programming approaches (Lamghari and Dimitrakopoulos 2012; Lamghari et al. 2014; Goodfellow and Dimitrakopoulos 2016; Montiel and Dimitrakopoulos 2017; 2018; Both and Dimitrakopoulos 2020; Paithankar et al. 2020).

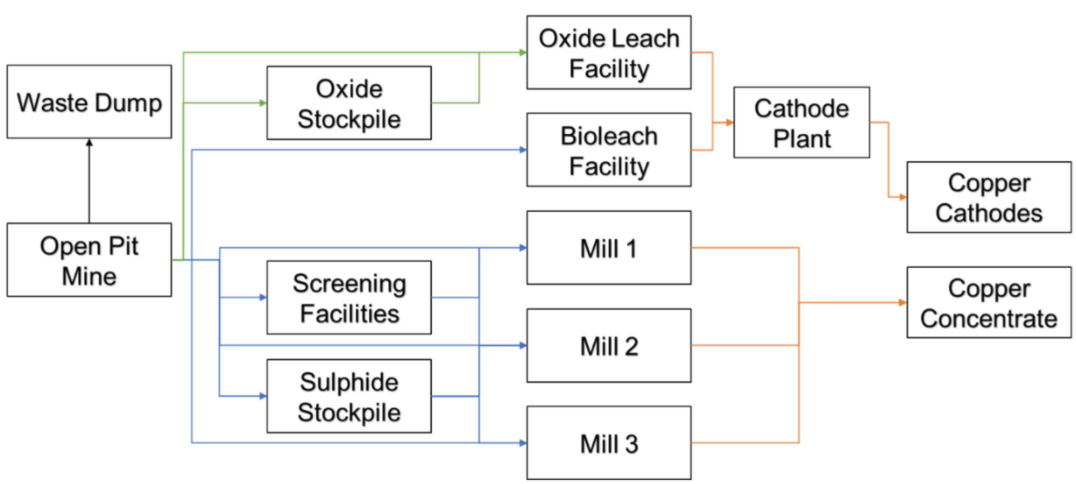


Figure 4.1. An example of an open-pit mining complex and material flow diagram (black: all materials, blue: sulphide ore, green: oxide ore and orange: recovered products).

Traditionally, production schedules are optimized beginning with the long-term schedule and then progressively scheduling shorter timescales to provide operational plans (Gershon 1983). The yearly long-term production schedule is optimized to maximize net present value over the life of a mining complex. Then, short-term production scheduling is completed to meet long-term production targets, which are often more detailed and consider additional operational requirements scheduled over shorter timescales (days, weeks, or months). Short- and long-term production schedules are optimized independently and the interaction between schedules are not considered during each optimization step. This causes misalignment across timescales and often leads to unrealistic forecasts and/or decreased project value. The ability to meet targets across planning horizons or production timescales is impacted by several factors that are considered in this work. This includes material heterogeneity, equipment accessibility and the capacities of different mining complex components. Similar to the benefits of optimizing components simultaneously, there is expected to be considerable upside in optimizing across different planning timescales.

This work proposes a novel stochastic optimization framework that connects planning horizons by jointly optimizing the short- and long-term production schedules in mining complexes. By connecting production scheduling timescales, the framework is expected to align the short- and long-term production forecasts and provide operationally feasible plans. Accounting for schedule alignment in the optimization framework helps ensure that the long-term production schedule is more attainable at shorter timescales, if needed. For example, there may be a forecast of 12 Mt of ore to be mined and sent to the processing facility over a production year, however, when scheduling at shorter timescales it is not possible to feed the constrained processing facilities 1 Mt per month to meet this forecast. Although this may be managed in the short-term by stockpiling material or mining large quantities of waste, having the inability to access ore in the mines can lead to operational risk and requires additional rehandling, which may not be economic. Ignoring short-term scheduling impacts may negatively affect the ability to meet long-term production targets leading to lower cashflows and metal production. This risk stems from the material heterogeneity and grade variability in the areas to be mined making it critical to consider the impact of short-term scheduling when creating the long-term plan. A stochastic programming formulation is proposed alongside a new solution approach that applies reinforcement learning to encourage compliance between short- and long-term production schedules to mitigate risk. Reasons for integrating reinforcement learning with stochastic programming optimization approaches are discussed subsequently. Previous attempts to connect short- and long-term production schedules focus on individual parts of the production schedule. Dimitrakopoulos and Jewbali (2013) consider the effect of simulating high-density future grade control data (only available in the short-term) and updating the simulations of the material supply to consider its impact on the long-term production schedule. Nehring et al. (2010) minimize deviations from a predefined mill feed grade over shorter timescales, while producing the long-term production schedule in an underground mine. Each method partially addresses the effect of optimizing over different timescales, however, they do not enforce schedule alignment and have not been extended to mining complexes.

In this work, reinforcement learning is explored to approximate the impact of long-term scheduling decisions at shorter timescales. Reinforcement learning is an approach where a policy is learned to maximize cumulative reward or return (Sutton and Barto 2018). Through several interactions, a reinforcement learning agent (or decision maker) learns the association between situations, actions and the rewards (or punishments) received. Actions that lead to higher rewards are learned and repeated over time by adjusting the policy to maximize future return. Recently, reinforcement learning techniques have been applied to optimize short-term production schedules with a neural network policy (Levinson et al. 2023). The approach proposed in this manuscript aims to connect planning horizons by embedding a reinforcement learning agent into the optimization framework to rapidly assess the short-term impact of long-term decisions, while optimizing the long-term production schedule. A key advantage of applying reinforcement learning in this setting is that the learned policy makes informed short-term scheduling decisions based on past experiences. This eliminates the need to reoptimize the entire short-term production schedule and provides an effective measure of its performance. Combining reinforcement learning and stochastic programming provides an opportunity to understand the impact of production scheduling decisions across timescales, which has not been considered in previous works due to the large number of additional decisions that must be considered. A related work that integrates learning-based aspects into optimizing mining complexes uses supervised machine learning to predict the value of mining extraction sequences to improve optimization efficiency (Lamghari et al. 2022). However, the main limitations of this approach are: (1) once the machine learning model is trained the parameters are left fixed and do not continue to adapt throughout the optimization; and (2) the method does not consider the connection between and short- and long-term term production schedules. Reinforcement learning approaches address these limitations by updating the policy with recent experiences and in this work connect planning horizons.

Reinforcement learning approaches for optimizing short-term production scheduling decisions demonstrate the ability to learn a neural network policy for specific tasks. For example, an actor-critic reinforcement learning framework with Monte Carlo tree search learns to adapt the extraction sequence and destination of materials given new incoming production data (Kumar and Dimitrakopoulos 2021). In addition, a gradient based policy optimization approach learns a destination policy for feeding material to one of several downstream process plants, oxide leach, bioleach and waste dump destinations for real-time decision making (Paduraru and Dimitrakopoulos 2019). Stochastic programming and actor-critic reinforcement learning have also been combined to optimize short-term production schedules in mining complexes using situational information taken directly from the production schedule (Levinson et al. 2023). The neural network parameters are optimized to maximize return by continuously interacting with a reinforcement learning environment and the learned policies can be applied to new situations to make short-term decisions without reoptimizing.

Jointly considering short- and long-term schedules allows for additional components to be considered in a mining complex. These may not impact the long-term production schedule significantly but can provide alternative material routing and short-term benefits by capitalizing on profitable and efficient operating alternatives. Preconcentration via screening is integrated into the short-term optimization model to separate material based on its grade-by-size behaviour, which provides opportunities to reject waste prior to energy intensive crushing and grinding at the processing facilities (Bowman and Bearman 2014; Carrasco et al. 2016a; Fathollahzadeh et al. 2021; Levinson et al. 2023). The proposed optimization formulation herein uses simulated response factors to account for uncertainty in the grade-by-size behaviour and allows for preconcentration facilities to eliminate certain size fractions of material with lower metal concentration by diverting material with beneficial properties. This increases the metal concentration of material sent to the process plant and reduces energy demand per unit of metal produced.

The first contribution of this work is a novel stochastic mathematical programming formulation that connects short- and long-term scheduling to align schedules of different timescales. Second, an optimization framework is developed that applies a data driven reinforcement learning approach for optimizing the short-term production scheduling decisions within a simulated annealing solution approach. The reinforcement learning agent interacts with a mining complex environment and continually learns from observed short-term production scheduling data and information related to the uncertainty and local variability of material grades to make informed short-term production scheduling decisions. The simulated annealing algorithm modifies long-term production scheduling decisions and the embedded reinforcement learning agent is applied to evaluate the impact on the short-term production schedule, eliminating the need reoptimize the entire short-term schedule by utilizing past learnings. Simulated annealing and reinforcement learning are applied together to optimize the stochastic programming formulation in an effective manner.

In the following sections of this paper, a stochastic mathematical programming formulation is introduced to balance short- and long-term production scheduling objectives and connect schedules of different timescales. Then, an optimization framework is proposed that combines simulated annealing and reinforcement learning techniques to optimize the presented stochastic programming formulation. Next, the proposed optimization framework is tested in a large-scale copper mining complex where the production schedules are compared with a traditional approach that optimizes the short- and long-term production schedules independently. Lastly, conclusions and future work are presented.

4.2 Stochastic programming formulation for connecting short- and long-term production schedules in mining complexes

A non-linear two-stage stochastic programming formulation is outlined for jointly optimizing short- and long-term production schedules in mining complexes. Supply uncertainty is incorporated with a set of stochastic orebody simulations (Goovaerts 1997; Rossi and Deutsch 2014). The formulation considers aspects for simultaneously optimizing the long-term production schedule and integrates additional short-term production scheduling components including preconcentration to construct a unified model for jointly optimizing short- and long-term production schedules.

Mining complexes may be composed of open-pit mines, stockpiles, preconcentration facilities, processing facilities (leach pads and process plants) and waste dump facilities. Mines supply material. Stockpiles store materials for later production periods. Preconcentration facilities sort and separate materials with beneficial grade-by-size characteristics and processing facilities liberate and recover valuable material from waste to generate valuable products. Waste dumps are large landforms that are constructed for disposing of low-quality or gangue material. An example of the components and material flows in a mining complex are shown in Figure 4.1. Materials from the mines are extracted, transported, and transformed through a series of connected components to generate valuable products for delivery to customers and the market. The costs for moving and processing material are incurred at each location in the mining complex and the revenues are obtained by selling the refined products. In the proposed formulation the extraction sequence, destination policy, stockpiling and preconcentration decisions are optimized simultaneously to determine the short- and long-term production schedules.

Objective Function

The objective function of the stochastic programming formulation shown in equation (4.1) aims to maximize net present value, based on the cash flow generated by the short- and long-term production schedules, manage technical risk and ensure production schedules of different timescales are achievable. The long-term production schedule is scheduled for each production year over the life of the mining complex and the short-term production schedule is scheduled within the first production year for each production month, week, or day. The definitions and notations for the sets, parameters and the short- and long-term decision variables are presented in Tables 4.1-4.4.

$$\max \frac{1}{\|S\|}\Bigg[\underbrace{\sum_{s\in S}\sum_{\tau\in T^{\mathrm{ST}}}\sum_{i\in\mathcal{D}}\sum_{q\in Q} p_{q,i,\tau}\left(v_{q,i,\tau,s}^{\mathrm{coarse}} + v_{q,i,\tau,s}^{\mathrm{fine}}\right)}_{\text{Part I}} + \underbrace{\sum_{s\in S}\sum_{t\in T^{\mathrm{LT}}\backslash\{1\}}\sum_{i\in\mathcal{D}}\sum_{q\in Q} p_{q,i,t} v_{q,i,t,s}}_{\text{Part II}}$$

$$- \underbrace{\sum_{s\in S}\sum_{\tau\in T^{\mathrm{ST}}}\sum_{i\in\mathcal{M}\cup\mathcal{D}} c_{i,\tau}\left(w_{i,\tau,s}^{\mathrm{coarse}} + w_{i,\tau,s}^{\mathrm{fine}}\right)}_{\text{Part III}} - \underbrace{\sum_{s\in S}\sum_{t\in T\backslash\{1\}}\sum_{i\in\mathcal{M}\cup\mathcal{D}} c_{i,t} w_{i,t,s}}_{\text{Part IV}}$$

$$- \underbrace{\sum_{s\in S}\sum_{t\in T}\sum_{i\in\mathcal{M}\cup\mathcal{D}} \left(\mathrm{pc}_{i,t}^{\mathrm{LT},-}\delta_{i,t,s}^{\mathrm{LT},-} + \mathrm{pc}_{i,t}^{\mathrm{LT}+}\delta_{i,t,s}^{\mathrm{LT},+}\right)}_{\text{Part V}}$$

$$- \underbrace{\sum_{s\in S}\sum_{\tau\in T^{\mathrm{ST}}}\sum_{i\in\mathcal{M}\cup\mathcal{D}} \left(\mathrm{pc}_{i,\tau}^{\mathrm{ST},-}\delta_{i,\tau,s}^{\mathrm{ST},-} + \mathrm{pc}_{i,\tau}^{\mathrm{ST},+}\delta_{i,\tau,s}^{\mathrm{ST},+}\right)}_{\text{Part VI}} - \underbrace{\sum_{s\in S}\sum_{a\in A} \mathrm{pc}_{a}^{-}\delta_{a,s}^{-}}_{\text{Part VII}}\Bigg] \tag{4.1}$$

Part I of the objective function (4.1) maximizes the discounted short-term revenues that are obtained by producing valuable products and selling them to customers and the market, given the short-term production scheduling decisions. Part II maximizes the discounted long-term revenues for producing valuable products for the long-term production periods. Part III and IV minimize short- and long-term discounted production costs that are incurred at each location in the mining complex to produce the valuable products. Part V and VI minimize deviations from long- and short- term production targets, respectively. These two components manage the risk of violating capacity constraints at the various locations in the mining complex (Goodfellow and Dimitrakopoulos 2016). Finally, Part VII connects planning timescales by ensuring that production schedules are obtainable in both the short- and long-term forecasts. This is accomplished with a set of soft constraints that measure deviations between long-term forecasts and the short-term forecasts for select attributes, for example, metal production. Deviations are minimized in the objective function to promote alignment. The objective function is optimized subject to the following constraints.

Constraints

Accessibility and reserve constraints enforce geotechnical slope constraints and ensure that a block can be mined at most once in the short- and long-term production schedule:

$$x_{b,t} \leq \sum_{k=1}^{t} x_{b',k} \quad \forall b \in B_m, b' \in \mathbb{O}(b), t \in T^{\mathrm{ST}} \cup T^{\mathrm{LT}}. \tag{4.2}$$

$$\sum_{t \in T^{\mathrm{LT}}} x_{b,t} \leq 1 \quad \forall b \in \mathcal{B}_m. \tag{4.3}$$

$$\sum_{\tau \in T^{\mathrm{ST}}} x_{b,\tau} \leq 1 \quad \forall\, b \in \mathcal{B}_m\,. \tag{4.4}$$

Mass and metal constraints are included to calculate the mass in each group of mined material (4.5-4.7) and the metal quantity during a short- or long-term production period (4.8-4.10):

$$\lambda_{g,t,s} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \theta_{b,g,s} w_{b,s} x_{b,t} \quad \forall g \in \mathcal{G}, t \in T^{\mathrm{LT}}, s \in S. \tag{4.5}$$

$$\lambda_{g,i,\tau,s}^{\mathrm{coarse}} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \mathrm{MR}_{b,i,s} \theta_{b,g,s} w_{b,s} x_{b,\tau} \quad \forall g \in G, i \in \mathcal{D}, \tau \in T^{\mathrm{ST}}, s \in S. \tag{4.6}$$

$$\lambda_{g,i,\tau,s}^{\mathrm{fine}} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} (1 - \mathrm{MR}_{b,i,s}) \theta_{b,g,s} w_{b,s} x_{b,\tau} \quad \forall g \in G, i \in \mathcal{D}, \tau \in T^{\mathrm{ST}}, s \in S. \tag{4.7}$$

$$\gamma_{q,g,t,s} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \theta_{b,g,s} g_{b,q,s} w_{b,s} x_{b,t} \quad \forall q \in Q, g \in \mathcal{G}, t \in T^{\mathrm{LT}}, s \in S. \tag{4.8}$$

$$\gamma_{q,g,i,\tau,s}^{\mathrm{coarse}} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} (1 - \mathrm{MR}_{b,i,s} \mathrm{RF}_{b,q,i,s}) \theta_{b,g,s} w_{b,s} g_{b,q,s} x_{b,\tau}$$
$$\forall q \in Q, g \in G, i \in \mathcal{D}, \tau \in T^{\mathrm{ST}}, s \in S. \tag{4.9}$$

$$\gamma_{q,g,i,\tau,s}^{\mathrm{fine}} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \mathrm{MR}_{b,i,s} \theta_{b,g,s} w_{b,s} g_{b,q,s} \mathrm{RF}_{b,q,i,s} x_{b,\tau}$$
$$\forall q \in Q, g \in G, i \in \mathcal{D}, \tau \in T^{\mathrm{ST}}, s \in S. \tag{4.10}$$

Destination policy constraints (4.11) allow a group of materials to only be sent to one destination in each production period:

$$\sum_{j \in \mathcal{D}} \kappa_{g,j,t} = 1 \quad \forall g \in \mathcal{G}, t \in T^{\mathrm{ST}} \cup T^{\mathrm{LT}}. \tag{4.11}$$

Constraints (4.12) compute the total material extracted from each mine in a production period:

$$w_{m,t,s} = \sum_{b \in B_m} w_{b,s} x_{b,t} \quad \forall m \in \mathcal{M}, t \in T^{\mathrm{ST}} \cup T^{\mathrm{LT}}, s \in S. \tag{4.12}$$

Material flow constraints ensure that there is a mass balance between downstream components of the mines. Constraints (4.13-4.15) track the recovery of each material property across the mining complex. Likewise, constraints (4.16-4.18) track the material mass throughout the mining complex. Constraints (4.19-4.20) allow stockpiles to be partially reclaimed and ensure all materials entering a processing or preconcentration facility are sent to a subsequent destination during the same production period. Constraints (4.21-4.22) ensure that the coarse and fine material fractions are only sent to a single destination in a short-term production period:

$$\begin{aligned} v_{q,j,(t+1),s} = &\sum_{g \in G} r_{q,j,(t+1)} \gamma_{q,g,(t+1),s} k_{g,j,(t+1)} + \sum_{i \in \mathcal{I}(j) \cap \mathcal{S}} r_{q,j,(t+1)} v_{q,i,t,s} y_{i,j,(t+1),s} \\ &+ v_{q,j,t,s} \left(1 - \sum_{k \in \mathcal{O}(j)} y_{j,k,t,s} \right) \quad \forall q \in Q, j \in \mathcal{D}, t \in T, s \in S. \end{aligned} \tag{4.13}$$

$$\begin{aligned} v^{\mathrm{coarse}}_{q,j.(\tau+1),s} = &\sum_{g \in G} r_{q,j,(\tau+1)} \gamma^{\mathrm{coarse}}_{q,g,j,(\tau+1),s} \kappa_{g,j,(\tau+1)} + \sum_{i \in \mathcal{I}(j) \cap \mathcal{F}} r_{q,j,(\tau+1)} v^{\mathrm{coarse}}_{q,i,(\tau+1),s} \, z^{\mathrm{coarse}}_{i,j,(\tau+1)} \\ &+ \sum_{i \in \mathcal{I}(i) \cap \mathcal{S}} r_{q,j,(\tau+1)} v^{\mathrm{coarse}}_{q,i,\tau,s} y_{i,j,(\tau+1),s} + v^{\mathrm{coarse}}_{q,j.\tau,s} \left(1 - \sum_{k \in \mathcal{O}(j)} y_{j,k,\tau,s} \right) \\ &\forall q \in Q, j \in \mathcal{D}, \tau \in T^{\mathrm{ST}}, s \in S. \end{aligned} \tag{4.14}$$

$$v_{q,j.(\tau+1),s}^{\text{fine}} = \sum_{g \in G} r_{q,j,(\tau+1)} \gamma_{q,g,j,(\tau+1),s}^{\text{fine}} \kappa_{g,j,(\tau+1)}$$
$$+ \sum_{i \in \mathcal{I}(j) \cap \mathcal{F}} r_{q,j,(\tau+1)} v_{q,i,(\tau+1),s}^{\text{fine}} \, z_{i,j,(\tau+1)}^{\text{fine}}$$
$$+ \sum_{i \in \mathcal{I}(i) \cap \mathcal{S}} r_{q,j,(\tau+1)} v_{q,i,\tau,s}^{\text{fine}} y_{i,j,(\tau+1),s} + v_{q,j.\tau,s}^{\text{fine}} \left(1 - \sum_{k \in \mathcal{O}(j)} y_{j,k,\tau,s} \right)$$
$$\forall q \in Q, i \in \mathcal{D}, \tau \in T^{\text{ST}}, s \in S. \tag{4.15}$$

$$w_{j,(t+1),s} = \sum_{g \in G} \lambda_{g,(t+1),s} k_{g,j,(t+1)} + \sum_{i \in (\mathcal{I}(j) \cap \mathcal{S})} w_{i,t,s} y_{i,j,(t+1),s}$$
$$+ w_{j,t,s} \left(1 - \sum_{k \in \mathcal{O}(j)} y_{j,k,t,s} \right) \quad \forall j \in \mathcal{D}\,, t \in T^{\text{LT}}, s \in S. \tag{4.16}$$

$$w_{j,(\tau+1),s}^{\text{coarse}} = \sum_{g \in G} \lambda_{g,j,(\tau+1),s}^{\text{coarse}} \kappa_{g,j,(\tau+1)} + \sum_{i \in \mathcal{I}(j) \cap \mathcal{F}} w_{i,(\tau+1),s}^{\text{coarse}} z_{i,j,(\tau+1)}^{\text{coarse}}$$
$$+ \sum_{j \in \mathcal{I}(i) \cap \mathcal{S}} w_{i,\tau,s}^{\text{coarse}} y_{i,j,(\tau+1),s}$$
$$+ w_{j,\tau,s}^{\text{coarse}} \left(1 - \sum_{k \in \mathcal{O}(j)} y_{j,k,\tau,s} \right) \quad \forall j \in \mathcal{D}, \tau \in T^{\text{ST}}, s \in S. \tag{4.17}$$

$$w_{j,(\tau+1),s}^{\text{fine}} = \sum_{g \in G} \lambda_{g,j,(\tau+1),s}^{\text{fine}} \kappa_{g,j,(\tau+1)} + \sum_{i \in \mathcal{I}(j) \cap \mathcal{F}} w_{i,(\tau+1),s}^{\text{fine}} z_{i,j,(\tau+1)}^{\text{fine}}$$
$$+ \sum_{j \in \mathcal{I}(i) \cap \mathcal{S}} w_{i,\tau,s}^{\text{fine}} y_{i,j,(\tau+1),s}$$
$$+ w_{j,\tau,s}^{\text{fine}} \left(1 - \sum_{k \in \mathcal{O}(j)} y_{j,k,\tau,s} \right) \quad \forall j \in \mathcal{D}, \tau \in T^{\text{ST}}, s \in S\,. \tag{4.18}$$

$$\sum_{i \in \mathcal{O}(i)} y_{i,j,t,s} \leq 1 \quad \forall i \in \mathcal{S}, t \in T^{\text{ST}} \cup T^{\text{LT}}, s \in S. \tag{4.19}$$

$$\sum_{i \in \mathcal{O}(i)} y_{i,j,t,s} = 1 \quad \forall i \in \mathcal{P} \cup \mathcal{F}, t \in T^{\text{ST}} \cup T^{\text{LT}}, s \in S. \tag{4.20}$$

$$\sum_{j \in \mathcal{O}(i)} z_{i,j,\tau}^{\text{coarse}} = 1 \quad \forall i \in \mathcal{F}, \tau \in T^{\text{ST}}. \tag{4.21}$$

$$\sum_{j \in \mathcal{O}(i)} z_{i,j,\tau}^{\text{fine}} = 1 \quad \forall i \in \mathcal{F}, \tau \in T^{\text{ST}}. \tag{4.22}$$

Production constraints (4.23-4.26) manage the risk of deviating from production capacities for short- and long-term production schedules at each location in a mining complex:

$$w_{i,t,s} - \delta_{i,t,s}^{\text{LT},+} \leq U_{i,t} \quad \forall i \in \mathcal{M} \cup \mathcal{D}, t \in T^{\text{LT}}, s \in S. \tag{4.23}$$

$$w_{i,t,s} + \delta_{i,t,s}^{\text{LT},-} \geq L_{i,t} \quad \forall i \in \mathcal{M} \cup \mathcal{D}, t \in T^{\text{LT}}, s \in S. \tag{4.24}$$

$$w_{i,\tau,s} - \delta_{i,\tau,s}^{\text{ST},+} \leq U_{i,\tau} \quad \forall i \in \mathcal{M} \cup \mathcal{D}, \tau \in T^{\text{ST}}, s \in S. \tag{4.25}$$

$$w_{i,\tau,s} + \delta_{i,\tau,s}^{\text{ST},-} \geq L_{i,\tau} \quad \forall i \in \mathcal{M} \cup \mathcal{D}, \tau \in T^{\text{ST}}, s \in S. \tag{4.26}$$

Attribute constraints (4.27-4.28) calculate the quantity of attributes across the optimization model as functions of the material qualities and quantities at each location in the mining complex:

$$v_{a,s}^{\text{ST}} = f_a\left(v^{\text{coarse}}, v^{\text{fine}}, w^{\text{coarse}}, w^{\text{fine}}\right) \quad \forall a \in A, s \in S. \tag{4.27}$$

$$v_{a,s}^{\text{LT}} = f_a(v, w) \quad \forall a \in A, s \in S. \tag{4.28}$$

Connection constraints (4.29) ensure long-term forecasts are achievable at shorter timescales and quantify deviations between short- and long-term forecasts:

$$v_{a,s}^{\text{ST}} + \delta_{a,s}^{-} \geq v_{a,s}^{\text{LT}} \quad \forall a \in A, s \in S. \tag{4.29}$$

Non-negativity constraints (4.30) ensure non-negativity for continuous variables; indices are omitted for presentation:

$$\delta, \gamma, \lambda, v, w, y \geq 0. \tag{4.30}$$

Binary constraints (4.31) enforce binary decisions; indices are omitted for presentation:

$$x, z \in \{0,1\}. \tag{4.31}$$

Table 4.1. Sets used in the stochastic programming formulation.

Sets	Descriptions
$\mathcal{M}$	Set of mines.
$\mathcal{S}$	Set of stockpiles.
$\mathcal{F}$	Set of preconcentration facilities.
$\mathcal{P}$	Set of processing facilities.
$\mathcal{W}$	Set of waste facilities.
$\mathcal{D}$	Set of destinations $\mathcal{D} = \{\mathcal{S} \cup \mathcal{F} \cup \mathcal{P} \cup \mathcal{W}\}$.
$\mathcal{G}$	Set of material groups, indexed by g.
$\mathcal{I}(i)$	Set of locations that supply material to location $i \in \mathcal{D}$.
$\mathcal{O}(i)$	Set of locations that receive material from location $i \in \mathcal{M} \cup \mathcal{D}$.
B_m	Set of blocks in the block model for each mine $m \in \mathcal{M}$, indexed by b.
$\mathbb{O}(b)$	Set of overlying blocks for each block $b \in B_m$.
S	Set of stochastic simulations of the uncertain material properties, indexed by s.
Q	Set of material quality properties, indexed by q.
A	Set of attributes tracked across the optimization model, indexed by a.
T^{ST}	Set of short-term production scheduling time periods.
T^{LT}	Set of long-term production scheduling time periods.

Table 4.2. Parameters used in the stochastic programming formulation.

Parameters	Description
$w_{b,s}$	Mass of block $b \in B_m$ in scenario $s \in S$.
$g_{b,q,s}$	Grade quality attribute $q \in Q$ in block $b \in B_m$ in scenario $s \in S$.
$\mathrm{RF}_{b,q,i,s}$	Response factor for attribute $q \in Q$ in block $b \in B_m$ at preconcentration facility $i \in \mathcal{F}$ in scenario $s \in S$, representing the grade-by-size behavior.
$\mathrm{MR}_{b,i,s}$	Mass response factor for block $b \in B_m$ at preconcentration facility $i \in \mathcal{F}$ in scenario $s \in S$, representing the percentage of material passing a screen.
$\theta_{b,g,s}$	Group membership parameter. Takes a value of one if block $b \in B_m$ is in group $g \in \mathcal{G}$ in scenario $s \in S$; zero otherwise.
d	Economic discount rate.
rd	Risk discount rate, for deferring risk to later production periods.
$p_{q,i,t}$	Selling price per unit of product produced with attribute $q \in Q$ and recovered at location $i \in \mathcal{P}$ in period $t \in T^{\mathrm{LT}}$, $p_{q,i,t} = p_{q,i,1}/(1+d)^t$.
$p_{q,i,\tau}$	Discounted selling price per unit of product produced with attribute $q \in Q$ and recovered at location $i \in \mathcal{P}$ in period $\tau \in T^{\mathrm{ST}}$.
$c_{i,t}$	Discounted unit cost to extract, handle and process material at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T^{\mathrm{LT}}$, where $c_{i,t} = c_{i,1}/(1+d)^t$.
$c_{i,\tau}$	Unit cost to extract, handle and process material at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $\tau \in T^{\mathrm{ST}}$.
$r_{q,i,t}$	Recovery of property $q \in Q$ at destination $i \in \mathcal{D}$ in period $t \in T^{\mathrm{LT}} \cup T^{\mathrm{ST}}$.
$\mathrm{pc}_{i,\tau}^{\mathrm{ST},+}, \mathrm{pc}_{i,\tau}^{\mathrm{ST},-}$	Penalty cost for a positive (+) or negative (-) deviation from short-term capacity at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $\tau \in T^{\mathrm{ST}}$.

$\mathrm{pc}_{i,t}^{\mathrm{LT},+}, \mathrm{pc}_{i,t}^{\mathrm{LT},-}$	Penalty cost for a positive (+) or negative (-) deviation from short-term capacity at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T^{\mathrm{LT}}$, where $pc_{i,t}^{\mathrm{LT},+/-} = pc_{i,t}^{\mathrm{LT},+/-}/(1+\mathrm{rd})^t$.
pc_a^-	Penalty cost for negative deviations from long-term forecasts in the short-term production schedule for each modeled attribute $a \in A$.
$L_{i,t}$	Lower bound on capacity at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T^{\mathrm{LT}} \cup T^{\mathrm{ST}}$.
$U_{i,t}$	Upper bound on capacity at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T^{\mathrm{LT}} \cup T^{\mathrm{ST}}$.

Table 4.3. Short-term decision variables used in the stochastic programming formulation.

Short-term Decision Variables	Description
$x_{b,\tau}$	Binary variable for each block $b \in B_m$ and production period $\tau \in T^{\mathrm{ST}}$. Takes a value of one if block b is mined in period t; zero otherwise.
$\kappa_{g,j,\tau}$	Binary variable determines whether material mined in group $g \in G$ is sent to destination $j \in \mathcal{D}$ during period $\tau \in T^{\mathrm{ST}}$. Takes a value of one if group g is sent to destination j in period τ; zero otherwise.
$y_{i,j,\tau,s}$	Continuous variable for determining the fraction of material sent from location $i \in \mathcal{D}$ to destination $j \in \mathcal{O}(i)$ in period $\tau \in T^{\mathrm{ST}}$.
$z_{i,j,\tau}^{\mathrm{coarse}}, z_{i,j,\tau}^{\mathrm{fine}}$	Binary decision variable defining the destination of material separated at screening facility $i \in \mathcal{F}$ and sent to destination $j \in \mathcal{O}(i)$ in period $\tau \in T^{\mathrm{ST}}$. Takes a value of one if coarse or fine material is sent from screen i to destination j during periot τ; zero otherwise.
$\delta_{i,\tau,s}^{\mathrm{ST},+}, \delta_{i,\tau,s}^{\mathrm{ST},-}$	Recourse variables measures positive (+) and negative (-) deviations from short-term production capacities at location $i \in \mathcal{M} \cup \mathcal{D}$ during period $\tau \in T^{\mathrm{ST}}$ and under scenario $s \in S$.

$v_{q,i,\tau,s}^{coarse/fine}$	Continuous variable denoting the quantity of property $q \in Q$ supplied to location $i \in \mathcal{D}$ in period $\tau \in T^{\mathrm{ST}}$ and scenario $s \in S$ in the coarse and fine size fractions.
$v_{a,s}^{\mathrm{ST}}$	Continuous variable that measures attribute $a \in A$ in scenario $s \in S$ considering short-term (ST) attributes.
$w_{i,\tau,s}^{\mathrm{coarse/fine}}$	Continuous variable denoting the mass of material received or mined at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $\tau \in T^{ST}$ and scenario $s \in S$ for both coarse and fine material.
$\lambda_{g,i,\tau,s}^{\mathrm{coarse/fine}}$	Continuous variable denoting the mass of the coarse and fine material in group $g \in G$ that can be sent to location $i \in \mathcal{D}$ in period $\tau \in T^{\mathrm{ST}}$ and scenario $s \in S$.
$\gamma_{q,g,i,\tau,s}^{\mathrm{coarse/fine}}$	Continuous variable denoting the quantity of property $q \in Q$ in the coarse and fine material in group $g \in G$ that can be sent to location $i \in \mathcal{D}$ in period $\tau \in T^{\mathrm{ST}}$ and scenario $s \in S$.

Table 4.4. Long-term decision variables used in the stochastic programming formulation.

Long-term Decision Variables	Description
$x_{b,t}$	Binary variable for each block $b \in B_m$ and production period $t \in T^{\mathrm{LT}}$. Takes a value of one if block b is mined in period t; zero otherwise.
$\kappa_{g,j,t}$	Binary variable determines whether material mined in group $g \in G$ is sent to destination $j \in \mathcal{D}$ during period $t \in T^{\mathrm{LT}}$. Takes a value of one if group g is sent to destination j in period t; zero otherwise.
$y_{i,j,t,s}$	Continuous variable for determining the fraction of material sent from location $i \in \mathcal{D}$ to destination $j \in \mathcal{O}(i)$ in period $t \in T^{\mathrm{LT}}$.

$\delta_{i,t,s}^{\mathrm{LT},+}, \delta_{i,t,s}^{\mathrm{LT},-}$ Recourse variables measures positive (+) and negative (-) deviations from long-term production capacities at location $i \in \mathcal{M} \cup \mathcal{D}$ during period $t \in T^{\mathrm{LT}}$ and under scenario $s \in S$.

$v_{q,i,t,s}$ Continuous variable denoting the quantity of property $q \in Q$ at location $i \in \mathcal{D}$ in period $t \in T^{\mathrm{LT}}$ and scenario $s \in S$.

$v_{a,s}^{\mathrm{LT}}$ Continuous variable denoting that measures attribute $a \in A$ in scenario $s \in S$ considering long-term (LT) attributes.

$w_{i,t,s}$ Continuous variable denoting the mass of material received or mined at location $i \in \mathcal{D}$ in period $t \in T^{\mathrm{LT}}$ and scenario $s \in S$.

$\lambda_{g,t,s}$ Continuous variable denoting the mass of material in group $g \in G$ in period $t \in T^{\mathrm{LT}}$ and scenario $s \in S$.

$\gamma_{q,g,t,s}$ Continuous variable denoting the quantity of property $q \in Q$ in group $g \in G$ in period $t \in T^{\mathrm{LT}}$ and scenario $s \in S$.

$\delta_{a,s}^{-}$ Recourse variables measures negative deviations from long-term production forecasts given the resulting short-term scheduling forecast for attribute $a \in \mathcal{A}$ at location $i \in \mathcal{M} \cup \mathcal{D}$ in scenario $s \in S$.

The following section outlines the optimization framework used to jointly optimize the proposed stochastic mathematical programming framework that connects short- and long-term production scheduling in mining complexes.

4.3 Optimization framework for connecting short- and long-term planning horizons

This section presents an optimization framework that can be used to optimize the non-linear stochastic programming formulation introduced in Section 4.2. The framework utilizes simulated annealing and reinforcement learning to optimize short- and long-term production scheduling decisions. Specifically, an actor-critic reinforcement learning approach is employed to determine the short-term decision variables with a learned neural network policy (Levinson et al. 2023). For the long-term decision variables, a modified simulated annealing algorithm is applied (Goodfellow and Dimitrakopoulos 2016). A key contribution of the proposed framework is that it incorporates a data driven reinforcement learning approach into the long-term optimization framework. The long-term decisions are modified with various strategies within a simulated annealing approach, while reinforcement learning is utilized to approximate the impact of the short-term production schedule given any modifications made to the long-term schedule. Notably, this combination of reinforcement learning and simulating annealing provides a novel and innovative approach to connect planning horizons in mining complexes that has not been explored before.

Figure 4.2 provides an overview of the optimization framework. The stochastic simulations of the material properties and mining complex parameters are input into the optimization framework in part A. The simulations represent the uncertainty and variability in the material grades and the mining complex parameters define the impact of different decisions that affect the movement of materials within a mining complex. Part B employs a modified simulated annealing approach that optimizes the long-term production schedule of the mining complex. The modified simulated annealing approach uses a single temperature parameter to look up the correct temperature for each decision neighbourhood in the simulated annealing algorithm based on the learned cumulative probability distribution functions and the objective function improvement, see Goodfellow and Dimitrakopoulos (2016) for further details. Following any modifications made to the long-term production schedule, current information related to the long-term schedule is collected and passed to the short-term reinforcement learning approach in part C. The embedded and trained reinforcement learning approach then optimizes the short-term production schedule in part D to determine the revenues, costs and deviations obtained by making the short-term scheduling decisions. These optimized short-term decisions are used to calculate the short-term components of the objective function (4.1) in part E and passed back to the simulated annealing framework. This iterative process ensures that short- and long-term production scheduling objectives are considered jointly. After each set of update iterations (n_{update}) are completed, part F of the framework is used to learn from previously encountered experiences. Data collected over the optimizations is used to update the neural network policy prior to optimizing the short-term production schedule in part D. This provides opportunities to improve the policy as the solution space is explored. Further details regarding each part of the framework are provided subsequently. All algorithms discussed can be found in Appendix A.

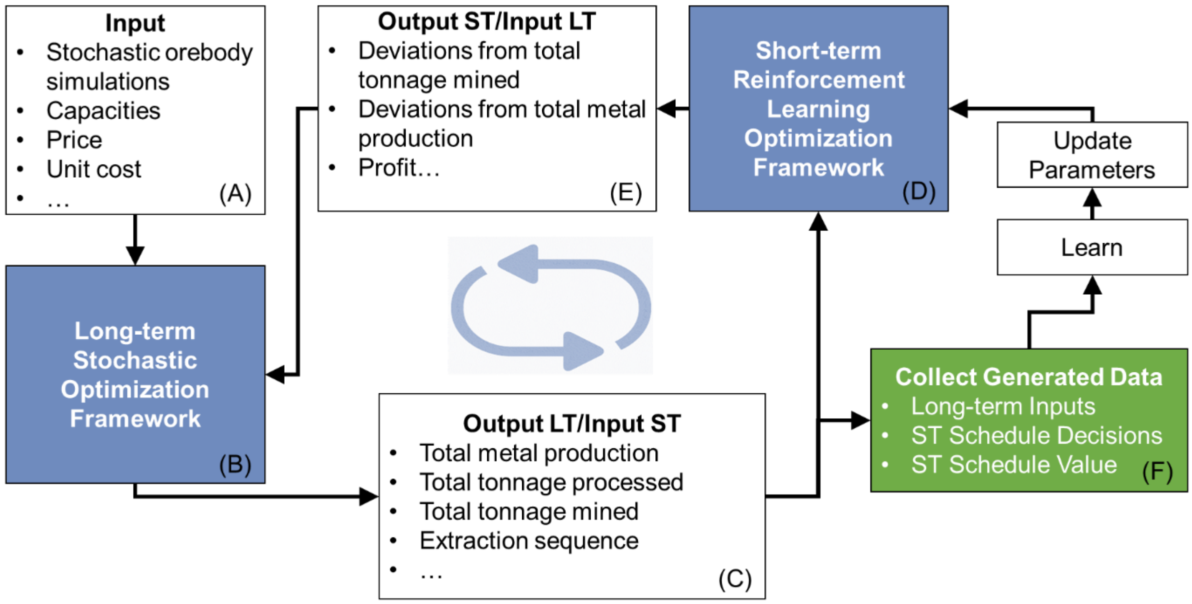


Figure 4.2. Optimization framework for connecting short (ST) and long (LT) production scheduling with simulated annealing and reinforcement learning.

4.3.1 Production scheduling with modified simulated annealing and an embedded reinforcement learning agent

The optimization framework requires a solution vector $\boldsymbol{\Phi} = \{\boldsymbol{x}, \boldsymbol{\kappa}, \boldsymbol{y}\}$ that stores the long-term production schedule decision variables where the extraction sequence vector $\boldsymbol{x}$ stores the variables $x_{b,t}$, the destination policy vector $\boldsymbol{\kappa}$ stores the variables $\kappa_{g,j,t}$ and the processing stream vector $\boldsymbol{y}$ stores processing stream variables $y_{i,j,t,s}, \forall m \in \mathcal{M}, b \in B_m, g \in G, i, j \in \mathcal{D}, t \in T^{\mathrm{LT}}, s \in S$. Algorithm 1 in Appendix A outlines the approach for optimizing the long-term production schedule. The modified simulated annealing algorithm begins with an initial feasible long-term solution vector $\boldsymbol{\Phi}$ and evaluates neighboring decisions by modifying the vector with a set of predefined perturbations (pb $\in$ PB). Each perturbation is constructed to satisfy the constraints outlined in the proposed stochastic programming formulation (Goodfellow and Dimitrakopoulos 2016; Montiel and Dimitrakopoulos 2017; Lamghari and Dimitrakopoulos 2020). However, in contrast to the approach taken by Goodfellow and Dimitrakopoulos (2016), short-term production scheduling decisions are determined using a reinforcement learning agent that obtains the short-term solution vector $\boldsymbol{\beta}$, given a short-term production scheduling policy π. The policy is learned with reinforcement learning and detailed in Section 4.3.2. The short-term solution vector $\boldsymbol{\beta} = \{\boldsymbol{x}^{\mathrm{ST}}, \boldsymbol{\kappa}^{\mathrm{ST}}, \boldsymbol{y}^{\mathrm{ST}}, \boldsymbol{z}^{\mathrm{ST}}\}$ stores the extraction sequence $x_{b,\tau}$, destination policy $\kappa_{g,j,\tau}$ and processing stream $y_{i,j,\tau,s}$ decision variables from the short-term schedule and introduces an additional vector $\boldsymbol{z}^{\mathrm{ST}}$ that stores the short-term specific $z_{i,j,\tau}^{\mathrm{coarse/fine}}$ decision variables for coarse preconcentration decisions, $\forall m \in \mathcal{M}, b \in B_m, g \in G, i, j \in \mathcal{D}, \tau \in T^{\mathrm{ST}}, s \in S$.

During each iteration of the simulated annealing algorithm, the modification to the long-term production schedule is communicated to the reinforcement learning framework. The reinforcement learning agent returns the short-term scheduling decision variables and the objective function is evaluated. If the value of the objective function increases after selecting an extraction sequence, destination policy or processing stream perturbation for the long-term production schedule and optimizing the short-term production schedule with the reinforcement learning agent the new solution is accepted. On the other hand, if the perturbation decreases the value of the objective function the modifications to the production schedules are accepted based on the acceptance probability of the selected perturbation following the distribution:

$$P(g(\mathbf{\Phi}, \boldsymbol{\beta}), g(\mathbf{\Phi}', \boldsymbol{\beta}'), \omega) = \begin{cases} 1 & \text{if } g(\mathbf{\Phi}', \boldsymbol{\beta}') > g(\mathbf{\Phi}, \boldsymbol{\beta}) \\ \exp\left(-\dfrac{\left(g(\boldsymbol{\Phi}', \boldsymbol{\beta}') - g(\boldsymbol{\Phi}, \boldsymbol{\beta})\right)}{\omega}\right) & \text{otherwise,} \end{cases} \tag{4.32}$$

where $g(\mathbf{\Phi}, \boldsymbol{\beta})$ and $g(\mathbf{\Phi}', \boldsymbol{\beta}')$ are the objective function value before and after a perturbation modifies the current short- and long-term solution vector. The modified simulated annealing algorithm retrieves ω the annealing temperature from a cumulative probability distribution of non-improving perturbations, for the perturbation selected, using a global annealing temperature parameter $\rho \in [0,1]$ (Goodfellow and Dimitrakopoulos 2016). The acceptance probability allows the algorithm to escape local optima by deteriorating the objective function. Furthermore, a cooling schedule for the global annealing temperature parameter is defined with a reduction factor $k \in$

$[0,1]$ and frequency parameter n that decreases ρ throughout the algorithm; the cooling schedule is applied to reduce the chance of accepting a deteriorating solution as the algorithm proceeds.

In the following section, the process for integrating the embedded actor-critic reinforcement learning agent is described along with the steps required to retrieve the short-term solution vector.

4.3.2 Embedding actor-critic reinforcement leaning agent

Reinforcement learning is used to optimize the short-term production schedule within the extraction limits of the first long-term production period ($t = 1$). The approach considers critical operational requirements often ignored in long-term production scheduling including mineability (bench and ramp access) and short-term operating alternatives that can improve efficiency and reduce costs, such as preconcentration. The short-term production schedule discretizes the areas to be mined in the first long-term production year into shorter timescales (i.e., months, weeks, days) providing increased granularity that can help identify and mitigate operational challenges, such as short-term material access and production bottlenecks. These challenges can have significant implications for meeting long-term production targets, particularly when the material supply is heterogeneous and uncertain. By incorporating short-term scheduling considerations into the optimization, the embedded reinforcement learning is expected to provide more realistic and robust production forecasts that consider operational impact.

A reinforcement learning setting is designed for optimizing short-term production scheduling decisions in mining complexes. An agent or learner interacts with a mining complex, describing the reinforcement learning environment, by determining short-term production scheduling decisions or actions over a series of timesteps $\tau \in T^{\mathrm{ST}}$. During each timestep, the agent observes the current state of the mining complex s_τ and takes an action a_τ by sampling a learned policy π. After taking the action, the environment responds by providing the next state $s_{\tau+1}$ and a reward $r_{\tau+1}$ that provides feedback on the agent's performance. This process is repeated to learn a policy for short-term production scheduling, by continuously interacting with the environment and finding actions that maximize the cumulative reward obtained by the agent. Those actions that increase the value of the short-term production schedule are preferred and the reinforcement learning policy is adapted to take those actions more frequently.

In the proposed optimization framework, the short-term production scheduling decision vector $\boldsymbol{\beta}$ are optimized with the reinforcement learning approach outlined in Levinson et al. (2023). Actor-critic reinforcement learning is used to learn a neural network policy (Mnih et al. 2016; Sutton and Barto 2018) for optimizing short-term production scheduling decisions in a mining complex. The policy learns to make short-term production scheduling decisions given the quality of material available, long-term forecasts and capacities at each location, which define the state of a mining complex. The short-term production schedule can then be efficiently assessed using the learned policy to retrieve $\boldsymbol{\beta}$ by applying Algorithm 2. This process is repeated during each iteration of the simulated annealing algorithm to assess the impact of modifying long-term production scheduling decisions at shorter timescales.

Actor-critic reinforcement learning requires learning the policy $\pi(a|s;\boldsymbol{\theta})$ and state value function $\hat{v}(s;\boldsymbol{\theta}_v)$ (Sutton and Barto 2018). The state value function estimates the value of state s under policy π and in this case provides the expected value of the short-term production schedule. The value function parameters are denoted by $\boldsymbol{\theta}_v$ and are shown as being separate from the policy function parameters $\boldsymbol{\theta}$, however, in practice the policy and state value function parameters can be shared.

The policy function modifies short-term extraction sequence decisions $\boldsymbol{x}^{\mathbf{ST}}$ in the short-term solution vector $\boldsymbol{\beta}$. This is accomplished by applying the policy $\pi(a|s;\boldsymbol{\theta})$ to map the state of the mining complex to the mining blocks extracted in each short-term production period $\tau \in T^{\mathrm{ST}}$. In this case, a specialized action space is used to transform the policy function output into actions designed for modifying short-term production scheduling decisions that directly ensure bench and ramp access in the mines. The number of mining access points that provide bench and ramp access in a mining complex are denoted by n_{ap}. The policy divides the actions into two parts for selecting extraction decisions at each access point AP_j , $j = 1, \ldots, n_{\mathrm{ap}}$. First, the mining direction is selected (east, west, north, south, or radial), a discrete action a_j^d, and second the quantity of material to be mined, a continuous action a_j^c, for all $j = 1, \ldots, n_{\mathrm{ap}}$. The blocks extracted by the agent are used to set the elements of the solution vector $\boldsymbol{x}^{\mathrm{ST}}$ to one if they are extracted in period τ and zero otherwise such that:

$$x_{b,\tau} = \begin{cases} 1 & \forall b \in N\left(\mathrm{AP}_j, a_j^d, a_j^c\right), j = 1, \dots, n \\ 0 & \text{otherwise,} \end{cases} \tag{4.33}$$

where $N\left(\mathrm{AP}_j, a_j^d, a_j^c\right)$ the set of blocks nearest to access point (AP_j) in the mining direction selected a_j^{d} and within the limits of the quantity extracted a_j^c. A couple of examples of the blocks mined considered the two action components at an access point are shown in a simplified 2-D example in Figure 4.3. A commercial solver (CLEX) is then used to determine the remaining short-term production scheduling decisions given the selected extraction sequence (Levinson et al. 2023). The reward function of the mining complex provides positive and negative feedback to the reinforcement learning agent by informing the agent of the impact of the actions selected. Each time an action is taken the experiences (s, a, r) obtained along each trajectory – the states, actions and rewards retrieved – are stored and later used for improving the short-term production scheduling policy.

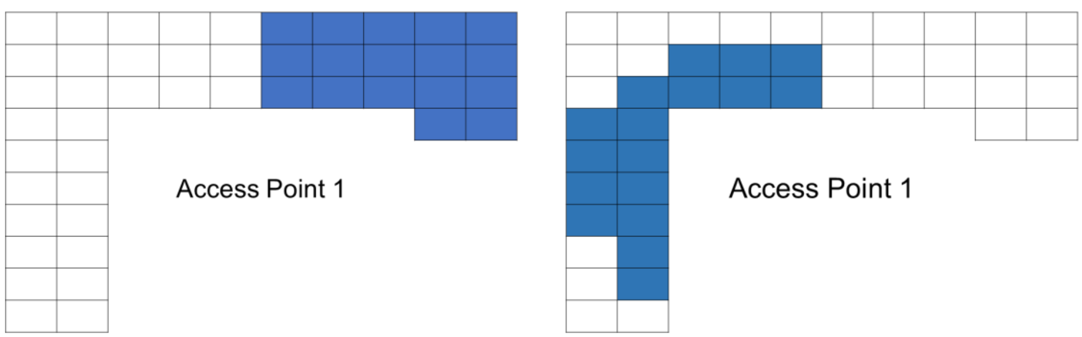


Figure 4.3. A simplified 2-D example of the parametrized policy using a continuous and discrete action space shows an east action selection (left) versus a west action selection (right) and the continuous variable is fixed to seventeen blocks in blue for both directions (Levinson et al. 2023).

4.3.3 Interactions between simulating annealing and reinforcement learning

The simulated annealing and reinforcement learning methods discussed in this section are applied to jointly optimize the short- and long-term production schedules in a mining complex. Algorithm 1 shows a modified simulating annealing approach that is used to explore different solutions by first modifying long-term production scheduling solution vector and assessing the short-term impact with an embedded reinforcement learning agent (Algorithm 2). The learned approximate decisions made with the reinforcement learning approach are used to evaluate the impact of short-term production scheduling decisions given the current solution vector for the long-term production schedule. The short-term production schedule and solution vector defined with the reinforcement learning policy affects the value of the objective function (4.1) in the stochastic programming formulation proposed in Section 4.2, which is computed in part B of the framework outlined in Figure 4.2. The resulting objective function value given the short- and long-term solution vectors is used to determine whether the modification to the long-term production schedule is accepted using the acceptance criterion in equation (4.32). Additionally, throughout the optimization, the neural network policy is improved by updating the neural network using the experiences obtained during short-term scheduling and applying a policy update with Algorithm 3 with an actor-critic approach (Sutton and Barto 2018; Paszke et al. 2019). The initial neural network policy and state value function are pretrained within the limits of a fixed long-term production schedule prior to the joint optimization of the short- and long-term schedule to obtain an initial policy and value function, as demonstrated in Levinson et al. (2023). The joint optimization is completed by applying Algorithm 4, which takes in the initial pre-trained neural network weights used to define the policy π and the initial long-term solution vector. Then, iteratively applies modified simulated annealing and updates the neural network policy incrementally throughout the optimization.

The proposed optimization framework for combining planning horizons in a mining complex with stochastic programming and reinforcement learning is applied at a copper mining complex in the next section.

4.4 Application in a copper mining complex

The open-pit copper mining complex considered is shown in Figure 4.1. A large open pit mine, with three primary materials sulfide ore, oxide ore and waste material, supplies raw materials to several downstream destinations. The downstream destinations include two stockpiles, a screening preconcentration facility, oxide leach, sulfide bioleach, three mills (process plants), and a waste dump facility. Extracted materials are transported from the mine to these destinations, separated, and beneficiated into two marketable products copper cathodes and copper concentrate for delivery to customers and the market. The uncertainty and local variability of the soluble copper, total copper and material response rank are quantified in the mine with a set of 15 equally probable geostatistical simulations (Albor Consuegra and Dimitrakopoulos 2009; Boucher and Dimitrakopoulos 2009; Montiel and Dimitrakopoulos 2017). There are 141,052 mining blocks available for extraction.

The destinations for sending material extracted from the mine are based on the material properties and types. Material types impact the processes required to liberate products from raw materials. For example, in this mining complex, sulfide materials are sent to either the sulfide stockpile, screening preconcentration facility, bioleach, one of three mills or the waste dump facility. Oxide materials can be sent the oxide stockpile, oxide leach or the waste facility and barren waste materials are sent directly to the waste facility. Each destination processes or stores material to take advantage of certain material properties and maximize the profitability of a mining complex. Stockpiles provide storage of material for processing in future time periods. Preconcentration destinations leverage materials with beneficial grade-by-size material properties (Carrasco et al. 2016a; Carrasco et al. 2016b). The preconcentration facilities provide an operating alternative that can be advantageous as it can reduce bottlenecks at the mill during operations (Burns and Grimes 1986; Levinson et al. 2023). The oxide leach pad facilities are used to recover oxide materials by applying acid leach solutions. Meanwhile, sulfide ores are recovered at the process plants where floatation takes place to produce copper concentrate. Lastly, bioleach facilities catalyze the oxidation of sulfide bearing minerals and provide a lower cost and lower recovery alternative to the process plant facilities, while extending the capacity for processing sulfide materials. Each processing destination has a fixed recovery parameter for the different products that can be recovered.

In the following sections, a comparison is made between two short- and long-term production schedules. The first long-term production schedule is optimized without considering short-term production scheduling aspects (Goodfellow and Dimitrakopoulos 2016). Then, similar to more standard approaches a separate short-term optimization is completed, within the limits of the first production year from the long-term schedule, using the optimization approach proposed in Levinson et al. (2023). These two schedules are used as the base case in the following sections. They show the resulting short- and long-term production schedules using a conventional approach that systematically optimizes across timescales starting with the long-term and then proceeding with the short-term. The risk profiles for these schedules are shown in grey in the following figures, where the P10, P50 and P90 represent the percentiles. The second short- and long-term schedules are optimized jointly using the proposed framework and compared to the base case production schedules. These risk profiles are shown in green in the following figures and consider the short-term integration within the optimization formulation (ST), as discussed in Section 4.2. The parameters used for the reinforcement learning approach are the same as those outlined in Section 3.5.2. In addition, the penalty costs used in the stochastic programming formulation relate to willingness to pay for a unit deviation from a constraint. These penalties were set to higher values for short-term objective components as the risk is more significant. Note that the penalty parameters used for managing risk in the optimization model are often manually adjusted to provide suitable risk profiles. Afterwards, the values are fixed for the comparisons to follow. The long-term schedules are scheduled over eight production years and the short-term production schedules are scheduled over twelve production months within the first production year. The two sets of short- and long-term production schedules are compared, and results discussed in the subsequent sections to address the advantages and limitations of the proposed formulation. The production schedules were optimized with an Intel® Core(TM) i7-8700 CPU and 32 GB of RAM and required 6.4-7.1 days to complete.

4.4.1 Short-term production schedule comparisons

The short-term production schedules shown in this section are used to compare the differences between jointly optimizing the short- and long-term schedule and separately optimizing the short- and long-term production schedules (the conventional base case). The resulting risk profiles for several critical components in the mining complex are discussed.

In Figure 4.4, the differences in the production at the mine (total material mined) in the short-term production schedule are shown. The short-term production forecasts are compared with the P-50 output of the long-term production schedule in the base case, which is divided by 12 months for comparisons to the short-term production schedule outputs. The short-term forecasts for the total material mined given the production schedule obtained with the joint optimization approach (green) demonstrates higher compliance with long-term forecasts when compared to the base case (grey). Obtaining the forecasted production levels for the total amount of material mined in the long-term forecast is critical in a mining complex, as the following production years are dependent on prior years being completed. Therefore, for the long-term schedule to be obtained those materials mined in the first year of production must be completely mined to ensure materials scheduled in the following year can be accessed. Figure 4.4 also shows the importance of integrating short-term operational decisions, such as, preconcentration via screening. Because the preconcentration decisions are included in the joint optimization model, the extraction sequence found with this approach prevents major deviations from the screen capacity, as seen in periods 1, 5 and 7 of the base case production schedule. However, if short-term scheduling aspects were not considered jointly these deviations would be undetectable in the long-term schedule, as the aggregated amount of material sent to the screen over the production year would not violate capacities in the long-term production schedule. Therefore, the higher fidelity model improves the decision-making process by accounting for short-term operational aspects in the joint optimization. This allows the optimizer to produce operationally realizable schedules, which account for violations that may be undetectable in conventional approaches until after optimizing the short-term production schedule. An advantage of optimizing across timescales is that it eliminates the need to adjust the long-term schedule to operational needs preventing these types of violations.

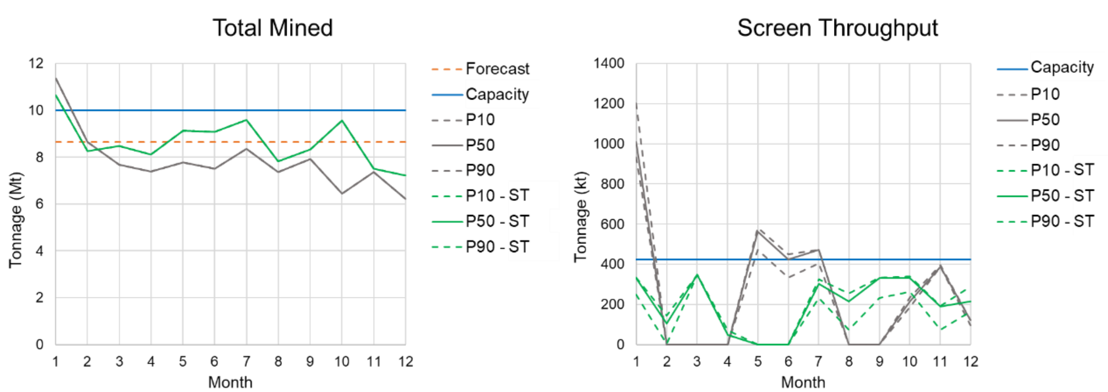


Figure 4.4. Risk profiles for total mined and preconcentration destination (screening facility) in the short-term production schedule.

Figure 4.5 and Figure 4.6 shows that the joint optimization utilizes the mill capacities more effectively by jointly optimizing the short- and long-term production schedule and overcomes the performance of the base case production schedule. The base case is unable to feed the mill consistently due to spatial operational limitations that are not considered until the short-term production schedule is defined. Although the required materials may be available within the first year of production based on the yearly forecast, the ability to access those materials in the correct proportions during each month within the production year is impossible with open-pit mining methods when operational accessibility is considered. This often occurs due to the heterogeneity and uncertainty of the material that relates to the location of high-quality materials in the mine. The joint optimization is equipped to adjust the long-term schedule to ensure the short-term schedule is obtainable at shorter timescales resulting in improved forecasts. Similar results are seen across processing destinations where the schedule that jointly considers the short and long-term production scheduling provides far more balanced and achievable production forecasts.

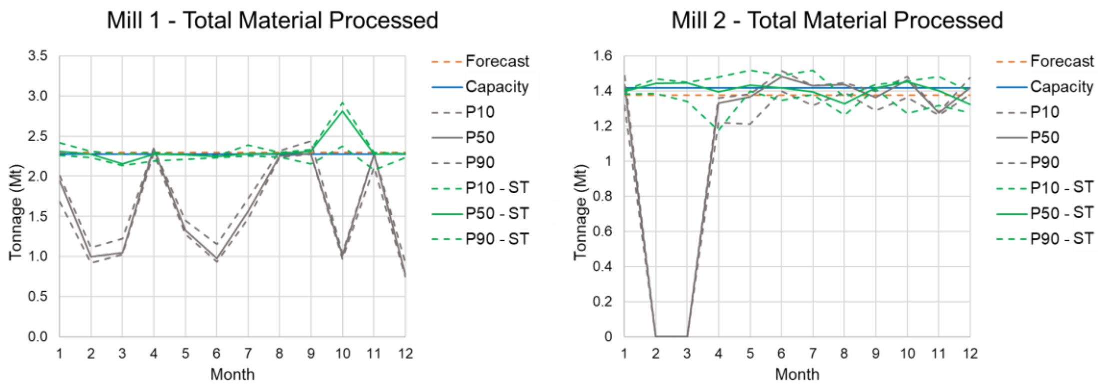


Figure 4.5. Risk profiles for total material processed at mill 1 and 2 in the short-term production schedule.

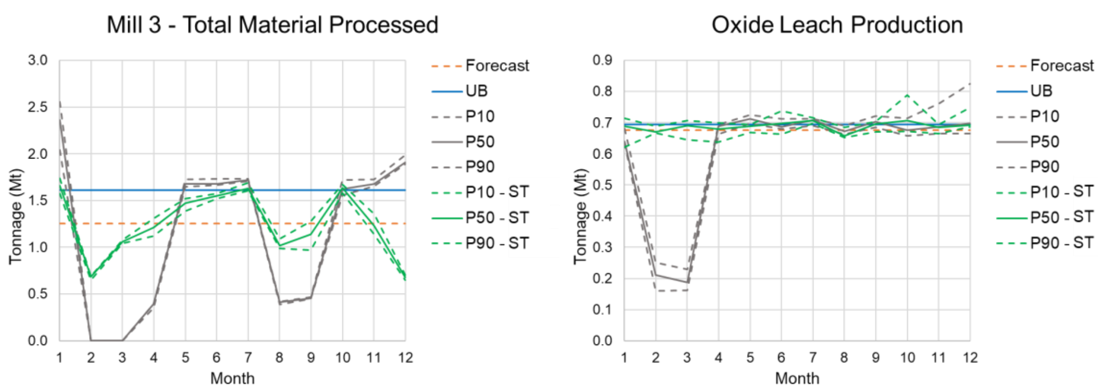


Figure 4.6. Risk profiles for total material processed at mill 3 and the oxide leach in the short-term production schedule.

Providing a stable feed and considering short-term operational constraints including accessibility in the joint optimization benefit the resulting financials and metal production. The joint optimization performs substantially closer to expected long-term forecasts based on the optimizers ability to adapt the long-term schedule to satisfy short-term requirements using the embedded reinforcement learning agent. Figure 4.7 show that the joint optimization of the short-term production schedule results in a 5.8% shortfall in the projected long-term cash flow whereas the base case schedule which optimizes them separately has a significant 26.8% shortfall on the forecasted cash flow over the twelve months. Jointly optimizing across timescales helps reduces deviations from long-term production cash flow in the first production year by 21%. Meanwhile, the cumulative metal production is similar in both schedules. This concludes the analysis of the short-term scheduling results. The impact of jointly optimizing the short- and long-term schedule on the long-term production schedule are considered in the following section.

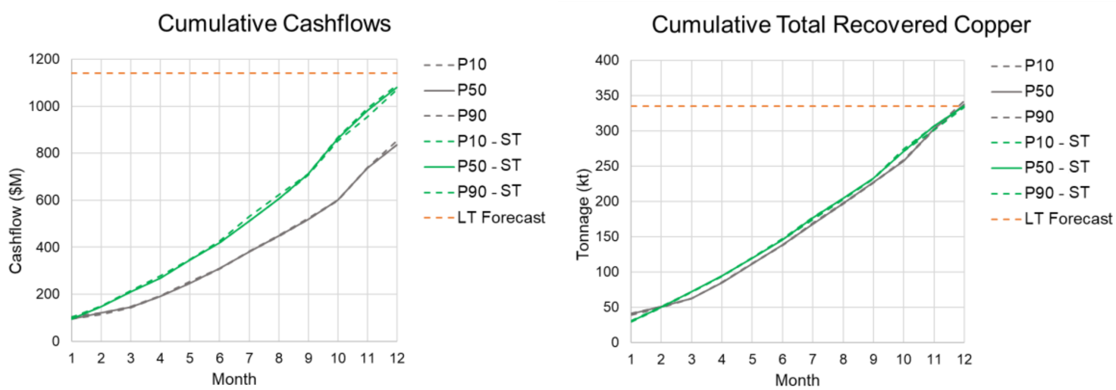


Figure 4.7. Risk profiles for cumulative cash flow and total recovered copper in the short-term production schedule.

4.4.2 Long-term production schedule comparison

In the previous section, substantial improvements were shown for the short-term production schedule in terms of production cash flow and plan compliance. This section now analyzes the two resulting long-term yearly schedules that are compared to observe the impact of considering jointly optimizing the short- and long-term production schedule (ST) and the conventional base case.

In Figure 4.8, the total material mined is shown on the left. The two production schedules have different mining rates from year-to-year; more material is mined in the first production year and less material in later production years when comparing the outcome of the joint optimization to the base case. The change in the long-term production schedule decreases the total amount of material mined by 5.9% over the life of mine, which lowers total mining costs.

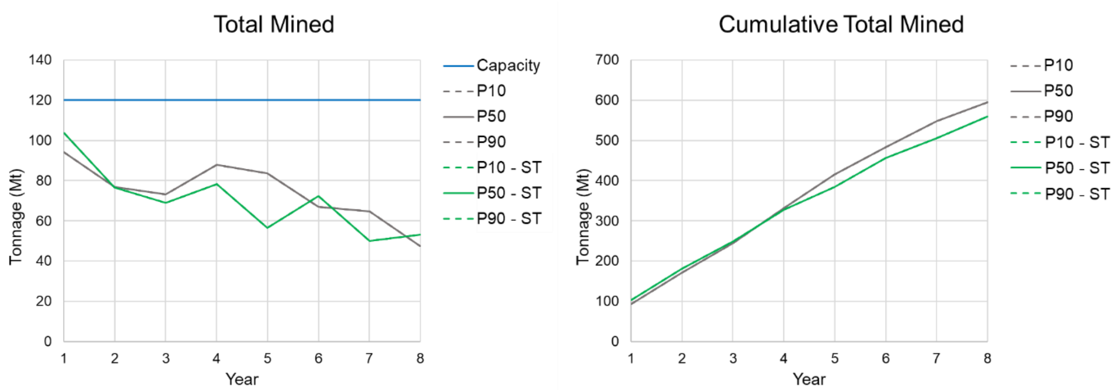


Figure 4.8. Risk profiles for total mined material and cumulative total mined material in the long-term production schedule.

In Figure 4.9, Mill 1 and Mill 2 are fully utilized in the first production period due to the ability to access and deliver material of the correct proportions. Similar to the discussion regarding the total material mined in the first year of production, the risk profiles for the mill feed in later period are also different as the schedule adapts the future mining areas to ensure short-term compliance. Similar results are shown at Mill 3 in Figure 4.10. The oxide leach risk profiles remain quite similar between the two production schedules. This is largely due to the oxide ore material being spread throughout the deposit at the surface.

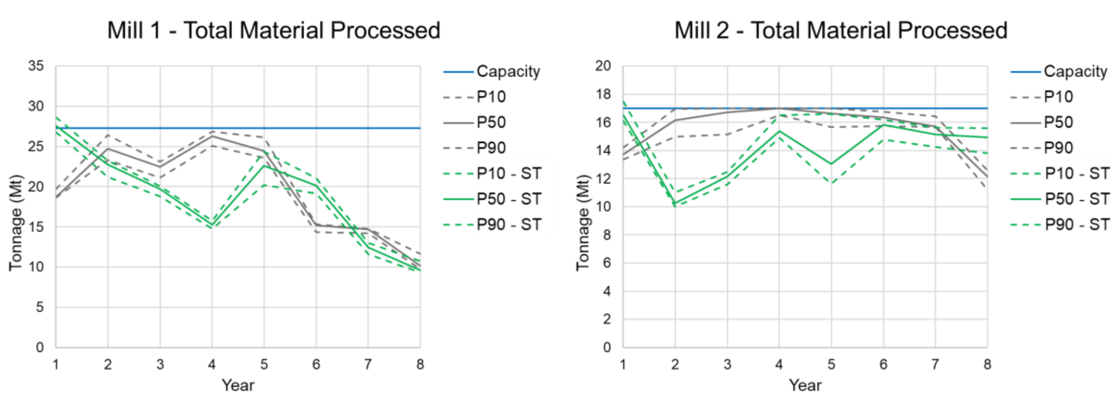


Figure 4.9. Risk profiles for total material processed at mill 1 and 2 in the long-term production schedule.

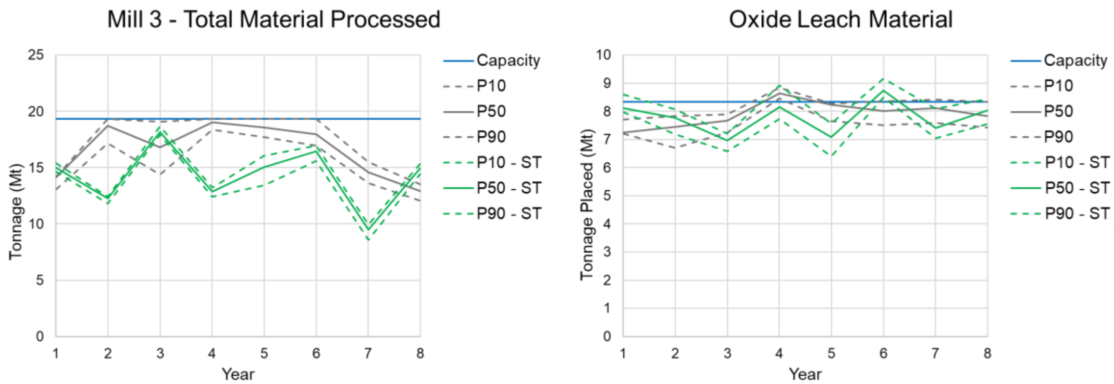


Figure 4.10. Risk profiles for total material processed at mill 3 and the oxide leach in the long-term production schedule.

In Figure 4.11, the net present value and cumulative copper production of the two long-term schedules are compared. After integrating the embedded reinforcement learning agent a 4% decrease in net present value is observed largely due to the need to adapt the schedule to the additional short-term operational requirements. By adapting the schedule to ensure a more feasible and attainable short-term schedule, a decrease in the long-term cash flow was found but, this forecast is realizable due to the short-term operational considerations included when scheduling the long-term production schedule. The results also indicate an 8% decrease in total copper production. These findings were anticipated as the joint optimization considers key operational aspects required for the successful extraction and processing of materials making forecasts substantially more realistic while attempting to sustain value. The major benefit of the proposed method is that the short-term production schedule is significantly more attainable due to the integration of the additional components in the objective function, which aim to enforce compliance across timescales. As shown in the short-term production schedule the resulting net present value of the base case is unlikely to be realized due to the inability to achieve production targets at shorter timescales. Therefore, the joint optimization is expected to improve performance and provide realizable production forecasts in the first production year, improving the reliability of the long-term forecasts.

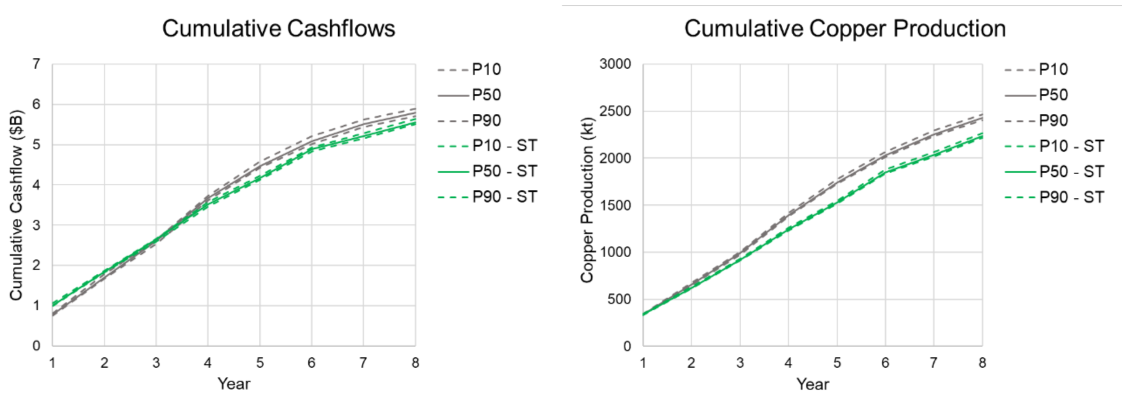


Figure 4.11. Risk profiles for cumulative cash flow and total recovered copper in the long-term production schedule.

Lastly, Figure 4.12 shows the comparison of the long-term schedule retrieved by jointly optimizing the short- and long-term production schedule compared to the base case production schedule. The figure highlights that there are large physical differences in the extraction sequence including the areas to be mined in each year and the extents of the pit. These changes are not limited to the first year of production but also later production years that are impacted due to small changes in the first production year that help satisfy short-term requirements. Both schedules remain mineable and only a cross-section of each long-term production schedule is shown.

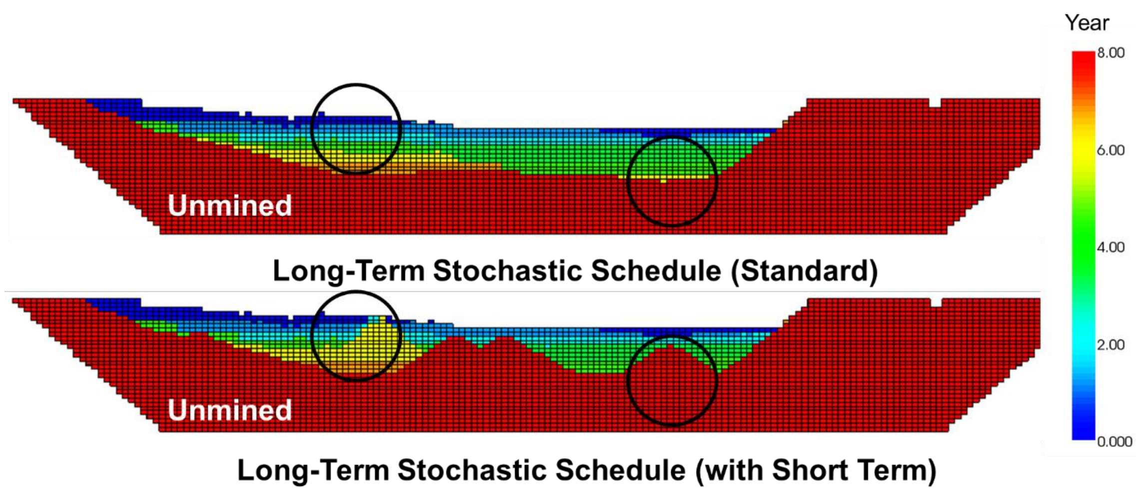


Figure 4.12. Cross-section of the long-term production schedule showing changes to the extraction sequence in early and late production periods; long-term schedule base case (top); long-term schedule with short-term integration (bottom).

Jointly optimizing the short- and long-term production schedule with stochastic programming and reinforcement learning provides an approach that can simultaneously optimize schedules of different timescales. The method eliminates the need to repeatedly adjust the long-term schedule to satisfy short-term operational requirements by evaluating them within the optimization framework. The results in the case study demonstrate that jointly considering short-term aspects in the optimization improves operational compliance and the resulting production forecasts at shorter timescales. This provides production forecasts that are informed by the impact of the interactions with short-term components and accounts for short-term risk related to the ability to operationally comply with the long-term plan. This has not been considered in past long-term production scheduling frameworks. In addition, the output schedules account for uncertainty and local variability of the material properties providing schedules that directly manage risk related to the material supply, while providing operational short-term plans. There are major advantages of incorporating additional operational synergies in the short-term schedule. These operational considerations may not be as advantageous in the long term but when planning at shorter time scales can reduce bottlenecks and improve the ability to meet production targets. For example, considering accessibility to material for the processing facilities and introducing preconcentration via screening in this case study led to improvements that provided a more stable production feed. Long-term production schedules that ignore these key operational impacts are unlikely to be realized and can mislead investors and operators based on their forecasted performance. The results shown in this case study are made possible by making approximate short-term scheduling decisions throughout the optimization with reinforcement learning. Reinforcement learning provides an effective method for evaluating short-term production schedules with a neural network policy; however, these decisions are approximate and significant work is still needed to ensure optimality and improve the overall optimization process. This work demonstrates an effective approach for jointly optimizing short- and long-term production schedules and is a starting point for future developments that can generate additional value and more accurate production forecasts by accounting for critical interactions across timescales.

4.5 Conclusions

A novel optimization framework is presented for connecting short- and long-term production scheduling in a mining complex under supply uncertainty. In order to account for the impacts of optimizing across different planning timescales, a stochastic mathematical programming formulation is defined. A modified simulating annealing approach is used to optimize the stochastic programming formulation, which utilizes an embedded actors-critic reinforcement learning approach to optimize the short-term production schedule efficiently with learned approximate decisions. Through the collection of experiences and updating of neural network policy, the reinforcement learning agent learns to make better short-term production scheduling decisions. The proposed framework eliminates the need to schedule different timescales separately and capitalizes on the synergies between them.

The application in a copper mining complex demonstrates the importance of planning across timescales using an integrative planning approach that ensures attainable short- and long-term production forecasts. Jointly optimizing the short- and long-term production schedule with reinforcement learning provides forecasts that are expected to be more accurate as short-term operational aspects including accessibility and coarse preconcentration decisions are considered. The long-term schedule adapts to these operational requirements and leads to significant improvements in long-term forecast compliance while maintaining value. The integrated approach decreases deviations from long-term operating cash flow by 21% by considering short-term aspects during the first production year. These changes to the short-term schedule impact future production periods in the long-term schedule decreasing the overall project value, however, when executed in operation the forecasts are expected to be more robust as the downstream destinations can be more effectively utilized.

Even though the first year of production is the most critical year for considering short-term operational aspects, the main reason for only considering the first production year in this work was to minimize additional computational overhead for evaluating the short-term schedule for additional production years. Future work should investigate parallel computing opportunities for distributing each production year to separate computer resources and scheduling the short-term during each long-term production year. In addition, further short-term components including equipment uncertainty and associated constraints should be considered. Lastly, this method is applied considering two timescales production years and months but the effects across shorter timescales i.e., weeks, days or hours should also be investigated. Furthermore, the developments presented herein can be extended to consider new incoming grade control information and other short-term production scheduling considerations to adapt the long-term production schedule to information that may not have been available at the time of planning helping minimize ore loss and dilution.

4.6 Appendix A: Pseudocode

Algorithm 1 Modified Simulated Annealing with Embedded Short-term Scheduling Reinforcement Learning Agent

Require:
- $\mathbf{\Phi}^g$ ▷ Global long-term solution vector
- ρ, k, n ▷ Acceptance probability, cooling factor, cooling schedule
- $i_{\max}$ ▷ Max number of iterations
- $\mathrm{CDF} = \{\mathrm{cdf_{pb}} : \mathrm{pb} \in \mathrm{PB}\}$ ▷ Distributions of non-improving perturbations

```
function SimulatedAnnealing(Φ^g)
    i ← 0                                        ▷ Iteration counter
    β^g ← ScheduleShortTerm(Φ^g)                 ▷ See Algorithm 2
    Φ, Φ' ← Φ^g                  ▷ Current and perturbed long-term solution vector
    β, β' ← β^g                  ▷ Current and perturbed short-term solution vector
    δ ← 0                        ▷ Initialize temporary annealing temperature
    while true do
        i ← i + 1                                ▷ Update iteration counter
        if i mod n == 0 then
            ρ = ρ * k            ▷ Update the acceptance probability parameter
        end if
        Φ', δ, pb = PerturbSolution(Φ, ρ)
        β' ← ScheduleShortTerm(Φ')               ▷ Retrieve ST solution vector
        r ← U[0, 1]                              ▷ Sample a random uniform number
        if P(g(Φ, β), g(Φ', β'), δ) ≥ r then     ▷ Eq. 32
            Φ ← Φ'                     ▷ Update the current long-term solution
            β ← β'                     ▷ Update the current short-term solution
        end if
        if g(Φ', β') < g(Φ, β) then
            Update cdf_pb with |g(Φ', β') − g(Φ, β)|
        end if
        if g(Φ', β') ≥ g(Φ^g, β^g) then
            Φ^g ← Φ'                   ▷ Update global best long-term solution
            β^g ← β'                   ▷ Update global best short-term solution
        end if
        if i == i_max then
            break          ▷ Terminate after max number of annealing iterations
        end if
    end while
    return Φ^g, β^g
end function
```

Algorithm 2 Retrieve short-term production schedule with the learned policy

Require:
  $\mathbf{\Phi}'$ ▷ Current long-term solution vector
  T^{ST} ▷ Set of short-term production periods
  $\boldsymbol{\theta}', \boldsymbol{\theta}'_v$ ▷ Current policy and value function parameters
  B ▷ Set of visited trajectories
  function SCHEDULESHORTTERM($\mathbf{\Phi}'$)
    Initialize $\mathscr{s}$ ▷ Current state of the mining complex
    Initialize $\mathscr{T}$ ▷ Empty trajectory set
    Initialize $\boldsymbol{\beta}'$ ▷ Short-term solution vector $\{\boldsymbol{x}^{\mathrm{ST}}\,\boldsymbol{\kappa}^{\mathrm{ST}},\boldsymbol{y}^{\mathrm{ST}},\boldsymbol{z}^{\mathrm{ST}}\}$
    for each $\tau \in T^{\mathrm{ST}}$ **do**
      $\mathscr{a} \sim \pi(\cdot|\mathscr{s};\boldsymbol{\theta}')$
      Determine the extraction sequence given $\mathscr{a}$ ▷ Eq. 33
      Optimize short-term production schedule and update $\boldsymbol{\beta}'$
      Observe new state $\mathscr{s}'$ and retrieve reward $\mathscr{r}$
      $\mathscr{T} \leftarrow \mathscr{T} \cup (\mathscr{s}, \mathscr{a}, \mathscr{r})$ ▷ Store experience
      $\mathscr{s} \leftarrow \mathscr{s}'$ ▷ Update state to next timestep
    end for
    $B \leftarrow B \cup \mathscr{T}$
    return $\boldsymbol{\beta}'$
  end function

Algorithm 3 Policy improvement for short-term production scheduling

Require:
  $\boldsymbol{\theta}', \boldsymbol{\theta}'_v$ ▷ Current policy and value function parameters
  B ▷ Set of visited trajectories
  function UPDATENETWORKS($\boldsymbol{\theta}'$, $\boldsymbol{\theta}'_v$)
    Initialize $d\theta, d\theta_v$ ▷ The policy and value function gradients
    $d\theta \leftarrow 0$ and $d\theta_v \leftarrow 0$ ▷ Reset gradients
    for each $\{\mathscr{s}_0, \mathscr{a}_0, \mathscr{r}_0, \ldots, \mathscr{s}_t, \mathscr{a}_t, \mathscr{r}_t\} \in B$ **do**
      for each $i \in \{t-1, ..., t_0\}$ **do**
        $G \leftarrow r_i + \gamma V(\mathscr{s}_{i+1};\theta_v)$ ▷ If $\mathscr{s}_{i+1}$ is terminal $V(\mathscr{s}_{i+1};\theta_v) = 0$
        //Accumulate gradients w.r.t. parameters
        $d\theta \leftarrow d\theta + \nabla_\theta \log \pi(a_i|\mathscr{s}_i;\theta)(G - V(\mathscr{s}_i;\theta_v))$
        $d\theta_v \leftarrow d\theta_v + \nabla_{\theta_v}(G - V(\mathscr{s}_i;\theta_v))^2$
      end for
    end for
    $\theta \leftarrow \theta + d\theta$ ▷ Update parameters θ
    $\theta_v \leftarrow \theta_v + d\theta_v$ ▷ Update parameters θ_v
    $B \leftarrow \emptyset$ ▷ Clear used trajectories
  end function

Algorithm 4 Jointly Optimizing the Short and Long-term Production Schedule

Require:
$\boldsymbol{\theta}, \boldsymbol{\theta}_v$ ▷ Initial pre-trained neural network parameters
$\boldsymbol{\Phi} = \{\mathbf{x}, \boldsymbol{\kappa}, \mathbf{y}\}$ ▷ Initial long-term solution vector
n_{update} ▷ Update frequency of neural network
n_{global} ▷ Max number of global iterations
procedure JointOptimization($\boldsymbol{\Phi}, \boldsymbol{\theta}, \boldsymbol{\theta}_v$)
  $\boldsymbol{\theta}' \leftarrow \boldsymbol{\theta}$, $\boldsymbol{\theta}'_v \leftarrow \boldsymbol{\theta}_v$ ▷ Initialize neural network parameters
  $\boldsymbol{\Phi}^g \leftarrow \boldsymbol{\Phi}$ ▷ Initialize the global best solution vector
  $i_{\text{global}} \leftarrow 0$ ▷ Global iteration counter
  while true **do**
    $i_{\text{global}} \leftarrow i_{\text{global}} + 1$ ▷ Update global iteration counter
    $\boldsymbol{\Phi}^g, \beta^g \leftarrow$ SimulatedAnnealing($\boldsymbol{\Phi}^g$) ▷ See Algorithm 1
    if $i_{\text{global}} \mod n_{\text{update}} == 0$ **then**
      $\boldsymbol{\theta}' \leftarrow$ UpdateNetworks($\boldsymbol{\theta}', \boldsymbol{\theta}'_v$) ▷ See Algorithm 3
    end if
    if $i_{\text{global}} == n_{\text{global}}$ **then**
      Save solution vector $\boldsymbol{\Phi}^g, \beta^g$
      break ▷ Terminate once max global iterations are reached
    end if
  end while
end procedure

4.7 Chapter discussion and next steps

Chapter 4 provides a stochastic mathematical programming framework for connecting short- and long-term production schedules by embedding a reinforcement learning agent into a metaheuristic solution approach. The method jointly optimizes schedules of different timescales leading to improved compliance between production schedules. Practical and attainable short-term production schedules are found that follow long-term production guidance. Capacities at processing facilities are further utilized as the optimization framework adapts long-term production scheduling decisions to short-term capabilities. This also ensures that long-term forecasts are achievable at shorter timescales by considering the impact of accessibility to the heterogeneous material. Materials to be extracted in a single production year are not necessarily available in the same proportions of ore and waste in each short-term production period leading to the potential for gaps in available ore, when considering the long-term production schedule alone. Uncertainty and variability in the material grades and the dissemination of high grades in the mineral deposits impacts ones ability to get the correct proportions of the desired material during shorter timescales. Optimizing different planning timescales together provides a finer level of detail that enhances production forecasts. This is accomplished by embedding the reinforcement learning agent proposed in Chapter 3 into a metaheuristic solution approach. Although several key components are optimized together in this framework, waste management and reclamation are ignored and require further development to integrate into the simultaneous stochastic optimization framework for mining complexes. Waste management and reclamation play an important role in production planning that improve environmental performance and ensure the mining complex is returned to a productive post mining state. These critical production scheduling considerations are integrated into the simultaneous stochastic optimization frameworks discussed in Chapter 5.

5 Simultaneous stochastic optimization of mining complexes: Integrating progressive reclamation and waste management with contextual bandits

Managing and preventing environmental impact in mining complexes is an integral part of long-term production planning. Sustainable long-term solutions are required to mitigate environmental damage and return the environment to a productive and beneficial post-mining land use. A simultaneous stochastic optimization framework for optimizing long-term production schedules in mining complexes is developed that integrates waste management and progressive reclamation. Uncertainty related to production of acid rock drainage is managed by simulating the geochemical properties of waste rock and blending materials of different qualities. A waste dump placement schedule is jointly optimized with the extraction sequence, destination policy, stockpiling and reclamation decisions in a single stochastic mathematical programming framework. The waste dump placement schedule includes a schedule for progressively reclaiming waste dump facilities in parallel with the production of valuable products, directly improving environmental performance. Additionally, the simultaneous stochastic optimization framework explores the use of contextual bandits to improve the metaheuristic solution approach. The framework is tested in a multi-mine copper-gold mining complex leading to improved environmental performance by significantly reducing the risk of ARD and progressively reclaiming the waste dump facilities.

5.1 Introduction

Waste management and reclamation are essential operational aspects to be considered when planning and optimizing the long-term production schedule of a mining complex (McHaina 2001). An integrative planning approach is desired to satisfy the diverse set of environmental constraints associated with operating a mining complex, while providing an economically viable long-term production schedule. For successful reclamation and rehabilitation, a progressive approach is taken to accelerate the reclamation of mining waste dump facilities, eliminate liabilities and improve environmental performance (Sawatsky et al. 2000; Bolan et al. 2017). Furthermore, strategic waste management practices that mitigate acid rock drainage (ARD) and prevent metal leaching are considered to meet water-quality standards and prevent large remediation costs (Price 2003).

Environmental management is vital for mitigating long-term complications that can develop during and after production ceases (Asif and Chen 2016). Directly considering these components during the optimization of the long-term production schedule can alleviate environmental risk and is expected to enhance the financial and environmental performance of a mining complex over its lifetime.

Depending on the processes required to recover valuable products that are sold to customers and the market, mining complexes may contain several mines, stockpiles, preconcentration, processing, and waste dump facilities (Pimentel et al. 2010). Research related to the simultaneous stochastic optimization of mining complexes aims to globally optimize all critical operational aspects in a single stochastic mathematical programming formulation (Goodfellow and Dimitrakopoulos 2016; Dimitrakopoulos and Lamghari 2022). This includes optimizing the extraction sequence, destination policy, capital investments, stockpiling, ore blending, processing, transportation and operating alternative decisions, while managing uncertainty related to quantity and quantity of the material in the ground (Montiel and Dimitrakopoulos 2015; 2017; Goodfellow and Dimitrakopoulos 2016; 2017; Del Castillo and Dimitrakopoulos 2019; Saliba and Dimitrakopoulos 2019; 2020; Levinson and Dimitrakopoulos 2020a; Paithankar et al. 2020; Brika et al. 2023). Decisions are optimized to maximize net present value, by enhancing the value of the products sold, and minimizing technical risk related to the uncertain material supply. However, the management, treatment and reclamation of mining waste dump facilities has received limited attention to date within these simultaneous stochastic optimization frameworks.

Waste management and rehabilitation scheduling decisions are typically considered following the optimization of the long-term production schedule (Li et al. 2013; Li et al. 2014; Vaziri et al. 2021; Vaziri et al. 2022), thus, advantageous synergies between the extraction sequence, destination policy and waste production schedule remain unseen. Fu et al. (2019) formulate a mixed integer programming model that jointly optimizes an open pit mine production schedule and waste dump placement schedule. Additionally, the encapsulation of potentially acid generating (PAG) waste is included to mitigate ARD, an important waste management consideration. There are several limitations to this integrative approach that are addressed in this work. First, the method ignores material uncertainty and variability of the material properties and uses a deterministic (average type) model of the material attributes. Second, PAG and non-acid generating (NAG) waste materials are classified based on the geochemical properties within a selective mining unit (a mining block), which ignores the potential to blend waste materials of different qualities to mitigate ARD. Lastly, the progressive reclamation of waste dump facilities is not considered within the optimization formulation. Levinson and Dimitrakopoulos (2018) globally optimize the long-term production schedule with waste management considerations in an open-pit gold mining complex under uncertainty. The simultaneous stochastic optimization framework manages risk of uncertain waste production by decreasing the production of PAG waste material. However, this approach does not provide an executable waste dump production schedule for material placement. Ben-Awuah and Askari-Nasab (2013) consider scheduling mining cuts and waste dumps jointly; however, this method ignores material uncertainty, aggregates mining blocks to reduce the number of decisions variables and does not produce waste dump production schedule. This manuscript aims to overcome the addressed limitations of previous works.

ARD and associated metal leaching have large impacts on water resources and are a serious concern for mining complexes producing wastes with sulfide bearing minerals (Nordstrom 2011). Following exposure to oxidizing conditions, sulfide bearing waste materials can produce sulfuric acid. Sulfuric acid is neutralized by contact with other minerals in the waste material, for example, carbonates. A recent field evaluation that monitors geochemical blending of PAG waste materials with NAG waste materials has demonstrated a method for permanent prevention of ARD under certain conditions (Day 2022). More common practices for mitigating ARD aim to eliminate or reduce the rate of oxidation with water covers, treatment methods that add chemicals to acidic water, and various other techniques (Skousen et al. 1998; Skousen et al. 2017). These conventional approaches can pose long-term risk. Treatment-based approaches extend long-term recurring costs that effect the economics of a mining complex. Meanwhile, containing acid forming material under water covers and constructed impoundments is high risk due to the potential of failures and the extended duration which the materials must remain covered (Grimalt et al. 1999; Price 2003; Kossoff et al. 2014; Thompson et al. 2020). Geochemical blending mixes waste with different geochemical characteristics to overcomes these limitations by preventing the formation of ARD at the source – the waste dump facilities – without the need for water covers. Blending can effectively manage the net producing ratio (NPR) or the ratio of neutralization potential (NP) to acid potential

(AP) of the mixed waste product at the final destination (Lawrence and Scheske 1997; Mehling et al. 1997).

Integrating the geochemical blending of waste rock materials into the optimization of mining complexes increases the complexity of placing waste material. The geochemical properties of waste are required to blend materials and manage the risk of ARD. Therefore, geostatistical simulation techniques are needed to quantify the quality and quantity of the relevant attributes within the mineral deposits in a mining complex (Goovaerts 1997; Boucher and Dimitrakopoulos 2009). After extraction, the geochemical attributes are tracked across the mining complex to blend waste materials of different qualities and generate a waste product that performs as if it is non-acid generating. A waste dump production schedule or sequence of material placement is determined by optimizing the mining complex to generate safe products for placement in waste dump facilities, which makes these decisions manageable in large-scale industry applications. Simultaneously optimizing the long-term production schedule with the waste dump placement schedule is critical for these reasons. This work develops a long-term stochastic optimization formulation that manages waste material blending to prevent ARD. Waste materials are mixed within waste dump cells to create a safe waste product by managing the NPR ratio. Considering the blending of waste materials to mitigate ARD can directly change the response of other production scheduling decisions as the schedule adapts to find waste with the correct geochemical properties, while still aiming to maximize net present value.

Another critical environmental management consideration related to planning and scheduling an operating mining complex are reclamation and closure requirements outlined by the local government and stakeholders of a region (Straker et al. 2020). Waste dump facilities cover large areas of a mining complex and must be returned to a natural and productive state to minimize negative environmental effects. This includes re-establishing tree cover and native species and ensuring landforms are returned to suitable environmental conditions for the region (Grant and Koch 2007; Koch 2007; Zipper et al. 2011). Progressive reclamation provides opportunities to monitor the results of reclamation decisions and reduce risk as requirements for closure are meet (Straker et al. 2020). Additionally, progressive reclamation allows for the direct placement of topsoil/overburden material in reclamation areas. Direct placement improves soil development at waste dump facilities leading to more productive reclamation of native species (Holmes 2001;

Mackenzie and Naeth 2010; Macdonald et al. 2015). Financial securities provided to regulators for reclamation requirements can also be returned based on the environmental work completed with a progressive reclamation approach that occurs in parallel with production. Understanding and integrating these critical planning components into the optimization of the long-term production schedule provides opportunities to eliminate environmental risk by planting and reclaiming segments of the waste dump facilities throughout production. This makes it crucial to consider in the simultaneous stochastic optimization of mining complexes.

This manuscript presents an innovative stochastic mathematical programming formulation and optimization framework that directly incorporates waste management and progressive reclamation considerations into the simultaneous stochastic optimization of a mining complex. Different from previous stochastic programming formulations, the model integrates waste management through the blending and scheduling of waste materials to mitigate ARD. In addition, the production schedule for placing topsoil/overburden material is included to accelerate progressive reclamation at the waste dump facilities to minimize long-term liabilities and satisfy regulatory requirements. Directly considering these aspects within the simultaneous stochastic optimization framework improves the overarching long-term plan but also increases the size of the stochastic mathematical program. Intelligent solvers are required to optimize these large-scale optimization models.

Contextual bandits are explored in this work for selecting heuristics within a metaheuristic solver to optimize the proposed stochastic programming formulation. Metaheuristics are advantageous for optimizing large-scale stochastic programming models in reasonable timeframes while handling non-linear constraints and have been widely applied in long-term production scheduling (Godoy 2003; Kumral and Dowd 2005; Ferland et al. 2007; Lamghari and Dimitrakopoulos 2012; Lamghari et al. 2014; Kumral 2013; Shishvan and Sattarvand 2015; Montiel and Dimitrakopoulos 2015; 2017; 2018; Goodfellow and Dimitrakopoulos 2016; 2017; Senécal and Dimitrakopoulos 2020; Brika et al. 2023). Goodfellow and Dimitrakopoulos (2016) adapt the acceptance criterion with a cumulative distributive function for each decision neighbourhood to improve the simulation annealing optimization framework. Instead of modifying the acceptance criterion, a different optimization strategy is considered in this work that utilizes learnings from a contextual bandit algorithm (Sutton and Barto 2018). The algorithm selects an action that it expects will lead to larger improvements in the objective function of a stochastic mathematical program, while exploring less visited actions with an ϵ-greedy policy. During optimization, the learner or agent adapts the actions selected given the context of the mining complex environment. Therefore, as the agent learns, the actions selected during the optimization are targeting changes to the production schedule that are expected to provide the largest improvements to the long-term production schedule.

In the next sections, the stochastic programming formulation is introduced followed by the optimization framework that utilizes contextual bandits. Subsequently, the proposed formulation and solution approach are tested with a case study in a large-scale mining complex. Lastly, conclusions and future work follow.

5.2 Simultaneous stochastic optimization of mining complexes with progressive reclamation and waste management

5.2.1 Stochastic mathematical programming optimization formulation

A two-stage non-linear stochastic programming model is outlined in this section (Birge and Louveaux 2011). The proposed formulation provides a simultaneous stochastic optimization model for jointly optimizing the extraction sequence, destination policy, stockpiling, waste dump production schedule, waste blending and progressive reclamation. The waste dump production schedule considers the placement of rock and overburden/topsoil material, while blending waste of different qualities to mitigate the risk of ARD. The objective function of the stochastic mathematical program aims to maximize net present value and minimize deviations from production targets. Figure 5.1 provides examples of the waste dump considerations in the proposed model. A set of equiprobable stochastic realizations of the key material attributes and the geochemical characteristics of interest are used to account for uncertainty and local variability of ore and waste materials (Boucher and Dimitrakopoulos 2009). The stochastic mathematical programming formulation is described next.

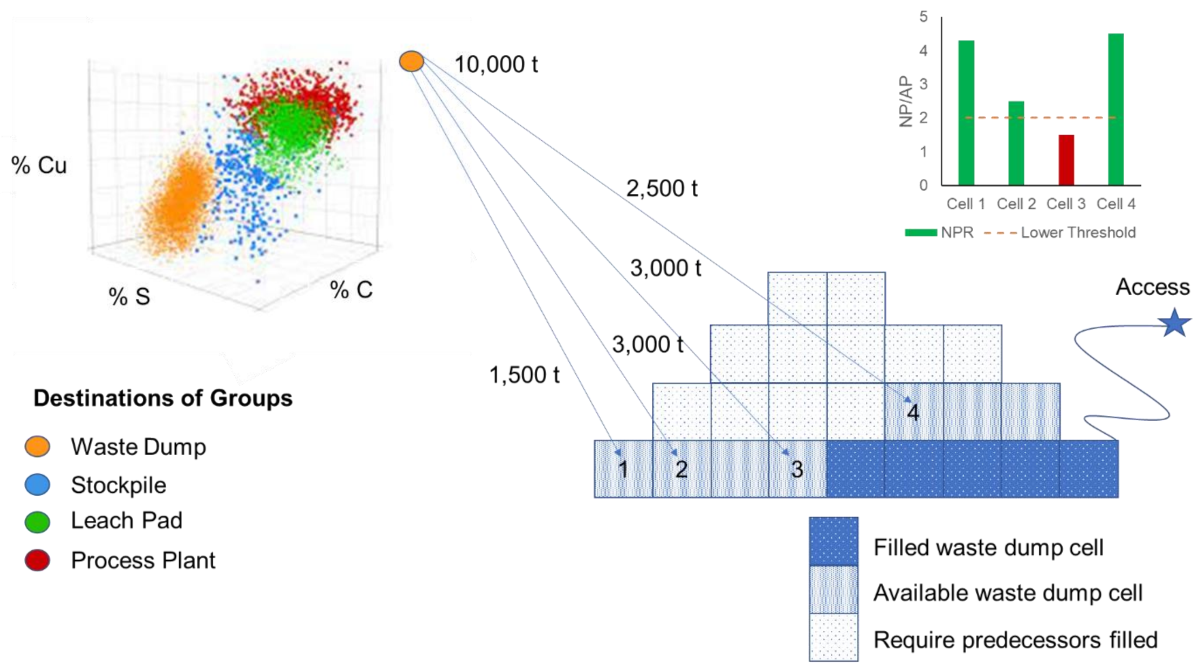


Figure 5.1. Blending strategy to manage NPR at waste dump facilities and precedence requirements in a simplified 2-D example of a waste dump model.

5.2.1.1 *Notations and definitions*

This section defines the notations and their definitions for the stochastic mathematical programming model. Table 5.1 and Table 5.2 include the sets and parameters. Table 5.3 includes the decision variables.

Table 5.1. Sets

Notations	Definitions
T	Set of long-term production periods to be scheduled, indexed by t.
S	Set of stochastically simulated block models of the pertinent material attributes, indexed by s.
A	Set of simulated material attributes in each block in the simulated block models, indexed by a.
$\mathcal{M}$	Set of mines.
$\mathcal{S}$	Set of stockpiles.

$\mathcal{P}$	Set of processing facilities.
$\mathcal{W}$	Set of waste dump facilities.
$\mathcal{D}$	Set of all mining complex destinations $\mathcal{D} = \mathcal{S} \cup \mathcal{P} \cup \mathcal{W}$.
C_w^{OB}	Set of waste dump cell locations at waste dump $w \in \mathcal{W}$ that accept topsoil/overburden material.
C_w^R	Set of waste dump cell locations at waste dump $w \in \mathcal{W}$ that accept rock material.
C_w	Set of all waste dump cell locations at waste dump $w \in \mathcal{W}$ where $C_w = C_w^{\mathrm{OB}} \cup C_w^R$, indexed by c.
G^{OB}	Set of material groups that are topsoil/overburden.
G^R	Set of material groups that are rock.
G	Set of all material groups such that $\mathcal{G} = \mathcal{G}^{\mathrm{OB}} \cup \mathcal{G}^R$, indexed by g.
$\mathcal{I}(i)$	Set of locations that accept material from location $i \in \mathcal{M} \cup \mathcal{D}$.
$\mathcal{O}(i)$	Set of locations that receive material from location $i \in \mathcal{M} \cup \mathcal{D}$.
B_m	Set of blocks in the block model for each mine $m \in \mathcal{M}$, indexed by b.
$\mathbb{O}(b)$	Set of overlying blocks for all blocks $b \in B_m$.
$\mathrm{PD}_{w,c}$	Set of predecessor cells to cell $c \in C_w$ that must be filled prior to placing material at cell c.

Table 5.2. Parameters

Notations	Definitions
$w_{b,s}$	Mass of block $b \in B_m$ in scenario $s \in S$.
$g_{a,b,s}$	Grade or quality of attribute $a \in A$ in block $b \in B_m$ and scenario $s \in S$.
$\mathrm{sf}_{b,s}$	Swell factor of block $b \in B_m$ in scenario $s \in S$ which represents the adjustment factor for converting bank volume to the loose material volume.

$\mathrm{sv}_{b,s}$	Specific volume (inverse of density) of the material in block $b \in B_m$ in scenario $s \in S$.
$\theta_{b,g,s}$	Block membership parameters set to one if block $b \in B_m$ is in group $g \in G$ in scenario $s \in S$, zero otherwise.
d	Economic discount rate.
rd	Risk discount rate for deferring risk to later production periods.
$p_{a,i,t}$	Selling price of a product with attribute $a \in A$ at location $i \in \mathcal{P}$ in period $t \in T, p_{a,i,t} = p_{a,i,1}/(1+d)^t$.
$r_{a,i,t}$	Recovery of product produced with attribute $a \in A$ at location $i \in \mathcal{D}$ in period $t \in T$.
$c_{i,t}$	Unit cost to extract, handle, or process material at location $i \in \mathcal{M} \cup \mathcal{D}$, $c_{i,t} = c_{i,1}/(1+d)^t$.
$\mathrm{rc}_{w,t}$	Reclamation cost per unit area at waste dump $w \in W$ in period $t \in T$.
ψ_c	Topsoil/overburden thickness requirement for cell $c \in C_w^{\mathrm{OB}}$.
$\mathrm{pc}_{i,t}^{+/-}$	Penalty cost for positive (+) or negative (-) deviations from the short-term capacity at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T$.
$\mathrm{pc}_{a,w,t}^{+/-}$	Penalty cost for positive (+) or negative (-) deviation from the upper or lower grade threshold of attribute $a \in A$ at waste dump location $w \in \mathcal{W}$ in period $t \in T$.
$\mathrm{pc}_{w,t}^{\mathrm{NPR}}$	Penalty cost for a unit deviation from the NPR threshold at waste dump location $w \in \mathcal{W}$.
$\mathrm{pc}_t^{\mathrm{Rec}}$	Penalty cost for a unit deviation from the reclamation target for the area reclaimed in production period $t \in T$.
$V_{w,c}$	Volume capacity for dumping material at waste dump $w \in \mathcal{W}$ in cell $c \in C_w$.
$L_{i,t}$	Lower bound on the capacity at location $i \in \mathcal{M} \cup \mathcal{D}$.
$U_{i,t}$	Upper bound on the capacity at location $i \in \mathcal{M} \cup \mathcal{D}$.
U_{Target}^{a}	Upper grade/quality threshold of metal leaching attribute $a \in A$ in the waste facility.

L^{a}_{Target}	Upper grade/quality threshold of metal leaching attribute $a \in A$ in the waste facility.
$L^{\text{NPR}}_{\text{Target}}$	Lower limit for NPR. Waste dump facility cells with an NPR ratio below this threshold are considered at risk of ARD. On the contrary, cells with an NPR ratio greater than this threshold are considered to be safe given the neutralization properties of the material within a waste dump cell.
NP	Indicating attribute of neutralization potential at the waste dump.
AP	Indicating attribute for acid production at the waste dump.
R_t	Annual reclamation target for waste dump facilities in period $t \in T$.

Table 5.3. Decision variables

Notations	Definitions
$x_{b,t}$	Binary variable set to one if block $b \in B_m$ is extracted in period $t \in T$, zero otherwise.
$y_{i,j,t}$	Continuous variable with range [0, 1] determines the fraction of material sent from stockpile location $i \in \mathcal{S}$ to destination $j \in \mathcal{O}(i)$ in period $t \in T$.
$z_{g,j,t}$	Binary variable set to one if material in group $g \in G$ is sent to destination $j \in \mathcal{D}$ in period $t \in T$, zero otherwise.
$\lambda_{g,t,s}$	Continuous variable denoting the mass of material extracted from group $g \in G$ in period $t \in T$ and scenario $s \in S$.
$\gamma_{a,g,t,s}$	Continuous variable denoting the mass of attribute $a \in A$ extracted from group $g \in G$ in period $t \in T$ and scenario $s \in S$.
$\overline{\text{mass}}_{i,t,s}$	Continuous variable denoting the mass of material passing through location $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T$ in scenario $s \in S$.
$\overline{\text{mass}}_{a,i,t,s}$	Continuous variable denoting the recovered quantity of attribute $a \in A$ at location $i \in \mathcal{D}$ in period $t \in T$ in scenario $s \in S$.
$\overline{\text{mass}}_{a,w,c,t,s}$	Continuous variable denoting the mass of attribute $a \in A$ at waste dump facility $w \in \mathcal{W}$ cell $c \in C_w$ in period $t \in T$ in scenario $s \in S$.
$\kappa_{w,c,t,s}$	Continuous variable in range [0,1] representing the percentage fill at waste dump $w \in \mathcal{W}$ cell $c \in C_w$ in period $t \in T$ and scenario $s \in S$.

$\alpha_{w,c,t,s}$	Binary variable denoting whether waste dump $w \in \mathcal{W}$ cell $c \in C_w$ is available for dumping material in period $t \in T$ and scenario $s \in S$.
$\overline{m}_{g,w,c,t,s}$	Continuous variable denoting the mass of material from group g sent to waste dump $w \in W$ to cell c given $(g, c) \in G^R \times C_w^R \cup G^{\mathrm{OB}} \times C_w^{\mathrm{OB}}$ in period $t \in T$, and scenario $s \in S$.
$\bar{v}_{g,w,t,s}$	Continuous variable denoting the volume of material from group $g \in G$ sent to waste dump $w \in \mathcal{W}$ in period $t \in T$ and scenario $s \in S$.
$g_{a,g,t,s}$	Continuous variable denoting the grade, quality, or concentration of attribute $a \in A$ within the material extracted from the mines in group $g \in G$, period $t \in T$ and scenario $s \in S$.
$\delta_{i,t,s}^{+/-}$	Recourse variable measuring positive (+) and negative (-) deviation from production targets at location $i \in \mathcal{M} \cup \mathcal{D}$ in period $t \in T$ and scenario $s \in S$.
$\delta_{a,w,c,t,s}^{+/-}$	Recourse variable denoting positive (+) and negative (-) deviations from material quality constraints for attribute $a \in A$ at waste dump $w \in \mathcal{W}$ at cell $c \in C_w$ in period $t \in T$ and scenario $s \in S$.
$\delta_{w,c,t,s}^{\mathrm{NPR}}$	Recourse variable denoting deviations from the NPR threshold in the blended product in waste dump $w \in \mathcal{W}$ cell $c \in C_w$ in period $t \in T$ and scenario $s \in S$.
$\delta_{t,s}^{\mathrm{Rec}}$	Recourse variable denoting deviations from the reclamation targets for each production period $t \in T$.
$\mathrm{sv}_{g,t,s}$	Continuous variable denoting the specific volume of material extracted from group $g \in G$ in period $t \in T$ and scenario $s \in S$.
$\mathrm{sf}_{g,t,s}$	Continuous variable denoting the swell factor of material extracted from group $g \in G$ in period $t \in T$ and scenario $s \in S$.

5.2.1.2 Objective function

The objective function of the stochastic programming model is shown in Eq. (5.1). Part I and II of the objective function maximize the revenues obtained from processing valuable attributes to sell to the market and minimize the costs to produce them. Part III of the objective function minimizes the reclamation costs per a unit area of reclaimed land at the waste facilities. This includes the placement, re-sloping, and planting costs for reclamation. Part IV minimizes deviations from long-term production targets related to the capacities of each location in the mining complex. Part V and VI minimize deviations from the blend of materials within the waste dump cell to prevent the production of ARD by managing the NPR ratio and metal leaching requirements. Part VII encourages progressive reclamation to place materials in locations that are prepared for reclamation. Maximizing the net present value by including the discounted revenues and costs associated with mining, processing and stockpiling material within the objective function has been considered in previous stochastic programming formulations while managing risk related to capacities (Montiel and Dimitrakopoulos 2015; Goodfellow and Dimitrakopoulos 2016). However, the contribution of this optimization model is that reclamation costs are directly considered in the objective function and the balancing of NPR and metal leaching material attributes are included to prevent ARD. Lastly, progressive reclamation is encouraged to mitigate liabilities and improve environmental performance of the mining complex.

$$\max \frac{1}{\|S\|}\left[\underbrace{\sum_{s\in S}\sum_{t\in T}\sum_{i\in\mathcal{P}}\sum_{a\in A} p_{a,i,t}\overline{\mathrm{mass}}_{a,i,t,s}}_{\text{Part I}} - \underbrace{\sum_{s\in S}\sum_{t\in T}\sum_{i\in\mathcal{M}\cup\mathcal{D}} c_{i,t}\overline{\mathrm{mass}}_{i,t,s}}_{\text{Part II}}\right.$$
$$- \underbrace{\sum_{s\in S}\sum_{t\in T}\sum_{w\in\mathcal{W}}\sum_{c\in C_w^{\mathrm{OB}}}\sum_{g\in G^{\mathrm{OB}}} (\mathrm{rc}_{w,t}/\psi_c)\,\mathrm{sf}_{g,t,s}\mathrm{sv}_{g,t,s}\overline{m}_{g,w,c,t,s}}_{\text{Part III}}$$
$$- \underbrace{\sum_{s\in S}\sum_{t\in T}\sum_{i\in\mathcal{M}\cup\mathcal{D}} \left(\mathrm{pc}_{i,t}^{-}\delta_{i,t,s}^{-} + \mathrm{pc}_{i,t}^{+}\delta_{i,t,s}^{+}\right)}_{\text{Part IV}} - \underbrace{\sum_{s\in S}\sum_{t\in T}\sum_{w\in\mathcal{W}}\sum_{c\in C_w} \mathrm{pc}_{w,t}^{\mathrm{NPR}}\delta_{w,c,t,s}^{\mathrm{NPR}}}_{\text{Part V}}$$
$$\left. - \underbrace{\sum_{s\in S}\sum_{t\in T}\sum_{w\in\mathcal{W}}\sum_{c\in C_w^{R}}\sum_{a\in A}\left(\mathrm{pc}_{a,w,t}^{-}\delta_{a,w,c,t,s}^{-} + \mathrm{pc}_{a,w,t}^{+}\delta_{a,w,c,t,s}^{+}\right)}_{\text{Part VI}} - \underbrace{\sum_{s\in S}\sum_{t\in T}\mathrm{pc}_{t}^{\mathrm{Rec}}\delta_{t,s}^{\mathrm{Rec}}}_{\text{Part VII}}\right] \tag{5.1}$$

The objective function is maximized subject to the following constraints.

5.2.1.3 Constraints

Block precedence and reserve constraints ensure that geotechnical slope constraints are enforced and that a single mining block can only be mined once during the long-term production schedule:

$$x_{b,t} \leq \sum_{k=1}^{t} x_{b',k} \quad \forall b \in B_m, b' \in \mathbb{O}(b), t \in T. \tag{5.2}$$

$$\sum_{t \in T} x_{b,t} \leq 1 \quad \forall b \in \mathcal{B}_m, m \in \mathcal{M}. \tag{5.3}$$

Mass and attribute tracking constraints for calculating material quantities throughout the mining complex. Constraints (5.4-5.5) compute the mass and attribute quantity extracted in each period.

$$\lambda_{g,t,s} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \theta_{b,g,s} w_{b,s} x_{b,t} \quad \forall g \in G, t \in T, s \in S. \tag{5.4}$$

$$\gamma_{a,g,t,s} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \theta_{b,g,s} w_{b,s} g_{a,b,s} x_{b,t} \quad \forall a \in A, g \in G, t \in T, s \in S. \tag{5.5}$$

Constraints (5.6) track the total amount of material mined at each mine in the mining complex:

$$\overline{\text{mass}}_{m,t,s} = \sum_{b \in B_m} w_{b,s} x_{b,t} \quad \forall m \in \mathcal{M}, t \in T, s \in S. \tag{5.6}$$

Constraints (5.7-5.8) track the mass and the recovered quantity of each attribute at the various destinations in mining complex. Constraints (5.9-5.10) ensure that only materials within a stockpile can be transported to a subsequent downstream destination:

$$\begin{aligned}\overline{\text{mass}}_{j,(t+1),s} &= \sum_{g \in G} \lambda_{g,t,s} z_{g,j,(t+1)} + \sum_{i \in \mathcal{I}(j) \cap \mathcal{S}} \overline{\text{mass}}_{i,t,s} y_{i,j,(t+1)} \\ &+ \overline{\text{mass}}_{i,t,s} \left(1 - \sum_{k \in \mathcal{O}(j)} y_{i,k,t}\right) \quad \forall j \in \mathcal{D}, t \in T, s \in S.\end{aligned} \tag{5.7}$$

$$\begin{aligned}\overline{\text{mass}}_{a,j,(t+1),s} &= \sum_{g \in G} r_{a,j,(t+1)} \gamma_{a,g,t,s} z_{g,j,(t+1)} + \sum_{i \in \mathcal{I}(j) \cap \mathcal{S}} r_{a,j,(t+1)} \overline{\text{mass}}_{a,i,t,s} y_{i,j,(t+1)} \\ &+ \overline{\text{mass}}_{a,j,t,s} \left(1 - \sum_{k \in \mathcal{O}(j)} y_{i,k,t}\right) \quad \forall a \in A, j \in \mathcal{D}, t \in T, s \in S.\end{aligned} \tag{5.8}$$

$$\sum_{i \in \mathcal{O}(i)} y_{i,j,t} \leq 1 \quad \forall i \in \mathcal{S}, t \in T. \tag{5.9}$$

$$\sum_{i \in \mathcal{O}(i)} y_{i,j,t} = 1 \quad \forall i \in \mathcal{D} \backslash \mathcal{S}, t \in T. \tag{5.10}$$

Constraints (5.11-5.13) calculate the swell factor, specific volume and volume of material sent from each group to the waste dump facility. The swell factor and specific volume of a group are calculated based on the input simulated block parameters and the extraction sequence selected such that:

$$\mathrm{sf}_{g,t,s} \lambda_{g,t,s} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \mathrm{sf}_{b,s} \theta_{b,g,s} w_{b,s} x_{b,t} \quad \forall g \in G, t \in T, s \in S. \tag{5.11}$$

$$\mathrm{sv}_{g,t,s} \lambda_{g,t,s} = \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \mathrm{sv}_{b,s} \theta_{b,g,s} w_{b,s} x_{b,t} \quad \forall g \in G, t \in T, s \in S. \tag{5.12}$$

$$\bar{v}_{g,w,t,s} = \mathrm{sf}_{g,t,s} \mathrm{sv}_{g,t,s} z_{w,g,t} \sum_{m \in \mathcal{M}} \sum_{b \in B_m} \theta_{b,g,s} w_{b,s} x_{b,t} \quad \forall g \in G, w \in \mathcal{W}, t \in T, s \in S. \tag{5.13}$$

Based on the material sent to the waste dump facilities the materials are distributed to different cells to blend materials of different qualities, as shown in Figure 5.1. Constraints (5.14-5.15) ensure that the volume of material extracted and sent to the waste dump facility are distributed to the waste dump cells that accept materials of specific types (rock and overburden):

$$\bar{v}_{g,w,t,s} - \mathrm{sf}_{g,t,s} \mathrm{sv}_{g,t,s} \sum_{c \in C_w^R} \bar{m}_{g,w,c,t,s} = 0 \quad \forall g \in G^R, w \in \mathcal{W}, t \in T, s \in S. \tag{5.14}$$

$$\bar{v}_{g,w,t,s} - \mathrm{sf}_{g,t,s} \mathrm{sv}_{g,t,s} \sum_{c \in C_w^{\mathrm{OB}}} \bar{m}_{g,w,c,t,s} = 0 \quad \forall g \in G^{\mathrm{OB}}, w \in \mathcal{W}, t \in T, s \in S. \tag{5.15}$$

Constraints (5.16-5.17) ensure that the volume of material placed in each cell does not violate the available volume of material within each waste dump facility cell:

$$\sum_{g\in G^{\mathrm{OB}}} \mathrm{sf}_{g,t,s}\mathrm{sv}_{g,t,s}\,\overline{m}_{g,w,c,t,s} \leq V_{w,c}\alpha_{w,c,t,s} \quad \forall w \in \mathcal{W}, c \in C_w^{\mathrm{OB}}, t \in T, s \in S. \tag{5.16}$$

$$\sum_{g\in G^{R}} \mathrm{sf}_{g,t,s}\mathrm{sv}_{g,t,s}\,\overline{m}_{g,w,c,t,s} \leq V_{w,c}\alpha_{w,c,t,s} \quad \forall w \in \mathcal{W}, c \in C_w^{R}, t \in T, s \in S. \tag{5.17}$$

Constraints (5.18-5.19) update the fill level of a waste dump cell:

$$\sum_{g\in G^{\mathrm{OB}}} \mathrm{sf}_{g,t,s}\mathrm{sv}_{g,t,s}\,\overline{m}_{g,w,c,t,s} = V_{w,c}\kappa_{w,c,t,s} \quad \forall w \in \mathcal{W}, c \in C_w^{\mathrm{OB}}, t \in T, s \in S. \tag{5.18}$$

$$\sum_{g\in G^{R}} \mathrm{sf}_{g,t,s}\mathrm{sv}_{g,t,s}\,\overline{m}_{g,w,c,t,s} = V_{w,c}\kappa_{w,c,t,s} \quad \forall w \in \mathcal{W}, c \in C_w^{R}, t \in T, s \in S. \tag{5.19}$$

Constraints (5.20) enforce precedence of the waste dump facility cells directly. The precedence constraints consider vertical and horizontal precedence relationships directly (a simplified example is shown in Figure 5.1):

$$\alpha_{w,c,t,s} \leq \kappa_{w,c',t,s} \quad \forall w \in \mathcal{W}, c \in C_w, c' \in \mathrm{PD}_{w,c}, t \in T, s \in S. \tag{5.20}$$

Constraints (5.21-5.23) compute the grade, quality or concentration of an attribute within a group which is used to determine the blended concentration of an attribute in the waste dump facility cells for overburden and rock material placement, respectively:

$$g_{a,g,t,s} = \gamma_{a,g,t,s}/\lambda_{g,t,s}. \tag{5.21}$$

$$\overline{\mathrm{mass}}_{a,w,c,t,s} = \sum_{g\in G^{\mathrm{OB}}} g_{a,g,t,s}\overline{m}_{g,w,c,t,s} \quad \forall a \in A, w \in \mathcal{W}, c \in C_w^{\mathrm{OB}}, t \in T, s \in S. \tag{5.22}$$

$$\overline{\text{mass}}_{a,w,c,t,s} = \sum_{g \in G^R} g_{a,g,t,s} \overline{m}_{g,w,c,t,s} \quad \forall a \in A, w \in \mathcal{W}, c \in C_w^R, t \in T, s \in S. \tag{5.23}$$

Constraints (5.24) manage grade blending to meet the NPR threshold target within each waste dump cell with a linear function f of the indicating geochemical attributes NP and AP. By enforcing, the following constraints it is possible to mitigate ARD in perpetuity by ensuring the ongoing prevention with a blended waste product that is non-acid generating:

$$\overline{\text{mass}}_{\text{NP},w,c,t,s} - \delta_{w,c,t,s}^{\text{NPR}} \geq L_{\text{Target}}^{\text{NPR}} f\left(\overline{\text{mass}}_{\text{AP},w,c,t,s}\right) \quad \forall w \in \mathcal{W}, c \in C_w, t \in T, s \in S. \tag{5.24}$$

Constraints (5.25-5.26) manage the risk of the concentration of geochemical attributes related to metal leaching exceeding the upper and lower concentration limits at the cells in a waste dump facility:

$$\overline{\text{mass}}_{\text{a},w,c,t,s} - \delta_{a,w,c,t,s}^{+} \leq U_{\text{Target}}^{a} \sum_{g \in G^R} \sum_{\tau=1}^{t} \overline{m}_{g,w,c,t,s}$$
$$\forall a \in A, w \in \mathcal{W}, c \in C_w^{\text{R}}, t \in T, s \in S. \tag{5.25}$$

$$\overline{\text{mass}}_{\text{a},w,c,t,s} + \delta_{a,w,c,t,s}^{-} \geq L_{\text{Target}}^{a} \sum_{g \in G^R} \sum_{\tau=1}^{t} \overline{m}_{g,w,c,t,s}$$
$$\forall a \in A, w \in \mathcal{W}, c \in C_w^{\text{R}}, t \in T, s \in S. \tag{5.26}$$

Constraints (5.27) aim manage the risk of meeting progressive reclamation targets for completed areas in each waste dump facilities:

$$\sum_{c \in c_w^{\text{OB}}} \kappa_{w,c,t,s} + \delta_{t,s}^{\text{Rec}} \geq R_t \quad \forall w \in \mathcal{W}, t \in T, s \in S. \tag{5.27}$$

Constraints (5.28-5.29) manage the risk of deviating from the capacity constraints at each location in the mining complex:

$$\overline{\mathrm{mass}}_{i,t,s} - \delta^{+}_{i,t,s} \leq U_{i,t} \quad \forall i \in \mathcal{M} \cup \mathcal{D}, t \in T, s \in S. \tag{5.28}$$

$$\overline{\mathrm{mass}}_{i,t,s} + \delta^{-}_{i,t,s} \geq L_{i,t} \quad \forall i \in \mathcal{M} \cup \mathcal{D}, t \in T, s \in S. \tag{5.29}$$

Constraints (5.30) only allow a group of materials to be sent to a single destination in each production period:

$$\sum_{j \in \mathcal{D}} \kappa_{g,j,t} = 1 \quad \forall g \in \mathcal{G}, t \in T^{\mathrm{ST}} \cup T^{\mathrm{LT}}. \tag{5.30}$$

Lastly, non-negativity and binary constraints (5.31-5.32) are enforced, indices omitted for clarity.

$$\bar{m}, y, \gamma, \delta, \kappa, \lambda \ \geq 0. \tag{5.31}$$

$$x, z, \alpha \in \{0,1\}. \tag{5.32}$$

The following section will outline the optimization approach for the proposed stochastic programming formulation.

5.2.2 Simulated annealing, contextual bandits and stochastic programming

The non-linear two-stage stochastic programming formulation is solved with a hybrid framework that utilizes a metaheuristic with embedded contextual bandits, and stochastic programming techniques. The metaheuristic applied is a simulated annealing-based approach that allows the model to handle non-linear objective function and constraints since the flow of materials are dependent on several key decision variables. Simulated annealing explores the neighbouring solutions by modifying the extraction sequence decision variables (Kirkpatrick et al. 1983; Goodfellow and Dimitrakopoulos 2016). The algorithm occasionally accepts solutions that decrease the objective function to escape local optima based on an acceptance criterion. The probability of accepting a deteriorating solution is adapted throughout the optimization by using a cooling schedule and annealing temperature parameter. The implementation of simulated annealing in the optimization of mining complexes has been applied in several previous works (Montiel and Dimitrakopoulos 2015; Goodfellow and Dimitrakopoulos 2016).

In this work, contextual bandits are integrated into the simulated annealing algorithm as a learning framework that aims to determine which perturbations (production schedule modifications) provide the largest improvements in the resulting objective function in Eq. (5.1) given some context of the modification. The actions and contexts that are used in this work are denoted by the sets K and C, respectively. Actions or perturbations that lead to larger improvements in the resulting production schedule of a mining complex in a given context are visited more frequently. This provides an effective way to improve the objective function by utilizing past learning to select more impactful perturbations during the optimization. Given the improvement and context of the optimization the resulting estimates of each action are updated for subsequent action selection. These estimates are referred to as estimates of the action-value function $Q(c, k)$ which provides the expected value of taking an action $k \in K$ in a given context $c \in C$. Within the simulated annealing framework, each action is selected with a ϵ-greedy policy which allows for exploration of different actions ($\epsilon \in [0,1]$ denotes the degree of exploration). The contextual bandit approach for selecting modifications to the production schedule is outlined in Algorithm 1 (Sutton and Barto 2018).

Algorithm 1 A contextual bandit algorithm

Require:

O ▷ Optimistic value above the maximum reward
α ▷ Learning rate
C ▷ Set of contexts
K ▷ Set of possible schedule modifying perturbations or actions
procedure ContextualBandit(O, α, C, K)
  Initialize the action-value function optimistically and the visit counts
  for $c \in C$ **do**
    for $k \in K$ **do**
      $Q(c, k) \leftarrow O$
      $N(c, k) \leftarrow 0$
    end for
  end for
  for $i = 1, .., N$ **do**
    c = get_mining_complex_context()

$$k = \begin{cases} \text{argmax}_k \text{Q(c,}\cdot\text{)}, & \text{with probability 1-}\epsilon \quad \text{(breaking ties with lowest visit count)} \\ \text{random action}, & \text{with probability } \epsilon \end{cases}$$

    Evaluate objective function after action is taken (and production schedule is modified)
    r = get_objective_function_change()
    $Q(c, k) \leftarrow Q(c, k) + \alpha(r - Q(c, k))$
    $N(c, k) \leftarrow N(c, k) + 1$
  end for
end procedure

Given that the optimization of mining complexes is a difficult learning task, the agent or learner must learn to track the best actions to take over time. This is because certain actions that perform better early in the optimization process (modifications to a larger number of decision variables) may not be the top performing modifications once the solution has been improved. For instance, smaller fine-tuning actions may become more beneficial at this time. Therefore, the goal is to learn a policy that maps a context to an action which is expected to improve the objective function by the greatest amount. The learned policy is constantly adapting to new information as the optimization algorithm proceeds. The algorithm proposed works as follows. First, the state-value function $Q(c,k)$ estimates are initialized to an optimistic value O. This encourages early exploration of all actions. Similarly, the visit count of each context-action pair $N(c,k)$ is initialized to zero. During the optimization process, each time a mining block is selected to modify the extraction sequence decision variables, the context is retrieved from the mining complex. Then, an action is taken with ϵ-greedy policy where the maximum action for the given context is taken with probability $1-\epsilon$. Otherwise, an action is selected uniformly at random. The action is applied to modify the extraction sequence decisions variables and the objective function value is updated after modifying these decisions. The difference between the current objective and the previous objective is calculated and used to update the action-value function. Lastly, the visit count for that action in this context is updated, which is used as a tie-breaking rule.

Similar to past work, the extraction sequence decisions are modified within the simulated annealing framework with a set of perturbations that ensure feasibility of the resulting extraction sequence (Goodfellow and Dimitrakopoulos 2016; Montiel and Dimitrakopoulos 2017; Lamghari and Dimitrakopoulos 2020). Then, the remaining downstream decision variables are optimized with the branch-and-cut algorithm implemented in CPLEX v20.1 given the temporarily fixed extraction sequence decisions to obtain the objective function value for a given iteration of the simulated annealing algorithm. This ensures that each extraction sequence evaluated within the simulated annealing algorithm has optimal downstream decisions. The proposed formulation and solution approach are tested at a large copper-gold mining complex in the following section.

5.3 Application in a large copper-gold mining complex

5.3.1 Overview of the copper-gold mining complex

The stochastic optimization formulation and solution approach are applied at a large-scale copper-gold mining complex to optimize the long-term production schedule. The mining complex under study has been operating for 10 years and has approximately 5 years of operational life remaining prior to closure where reclamation must be completed. The mining complex is presented in Figure 5.2. There are three open-pit mines, a run-of-mine stockpile, a processing facility, and three waste dump facilities that require mitigation of ARD and progressive reclamation. The open-pit mines supply materials to the processor within the mining complex to produce copper and gold products for delivery to customers and the market. Waste materials are separated into two material classes rock and overburden/topsoil. The geochemical characteristics of the materials in the deposit include simulations of the sulfur and total inorganic carbon content (assay samples are retrieved with an element scan with ICP-MS and aqua regia digestion and coulometry, respectively), which are used to calculate the NPR ratio in waste dump cells. In addition, the overburden/topsoil quantities are simulated to represent the uncertainty and variability in the quantity of material available for reclamation purposes (Strebelle and Cavelius 2014).

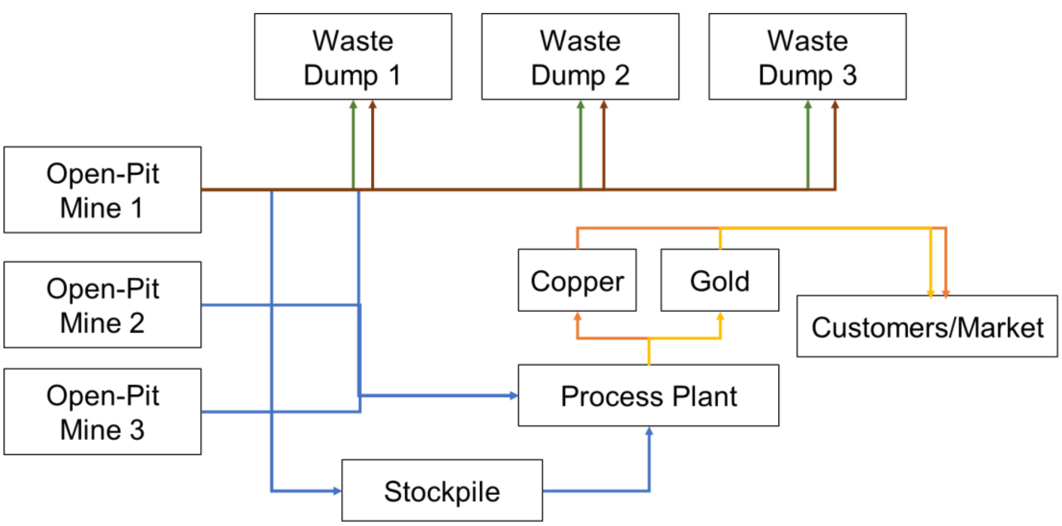


Figure 5.2. Mining complex with three open-pit mines, three waste dump facilities, stockpile, process plant that generates two valuable products.

The three open-pit mines that supply material within the mining complex have a joint mining capacity of 42 M tonnes per year and the process plant has a maximum capacity of 13 M tonnes with a recovery of 90% and 80.4% for copper and gold, respectively. Mining blocks are not aggregated and a set of 1.84 million blocks are considered for extraction with dimensions 10 m x 10 m x 16 m. Fifteen stochastically simulated orebody models with the pertinent attributes and geochemical properties are used as input to the optimization formulation to directly manage risk (Albor Consuegra and Dimitrakopoulos 2009; Montiel and Dimitrakopoulos 2015). The stochastic simulations of the orebody are simulated directly at block support and consider the correlations between attributes using an efficient method for quantifying the uncertainty and local variability of the material attributes in a large deposit (Boucher and Dimitrakopoulos 2009). The slope angle considered for mining block precedence is 47 degrees based on geotechnical studies completed by the operators of the mining complex.

The mining complex production schedule and waste dump placement schedule are optimized for the remaining five production years in this mining complex to minimize losses incurred to satisfy closure requirements. A waste dump model of the areas available for material placement in the waste dump facilities is constructed, similar to a block model, where the cell locations represent the volumes available for dumping materials in different areas. The ultimate goal of the long-term production schedule is to maximize net present value, while mitigating ARD in the waste dump facilities to prevent additional long-term monitoring costs. Additionally, a progressive reclamation strategy is implemented to start rehabilitating completed waste dump areas and return them to their original environmental state during production. This helps reduce liabilities and satisfy regulatory requirements. In addition, the progressive reclamation approach provides opportunities to improve reclamation practices if certain sections are unsuccessful. To begin reclamation in the waste dump facilities, underlying waste dump cells that are below reclamation cells must be completed to ensure the full utilization of areas available for waste disposal. A target progressive reclamation goal of 50 hectares (ha) is targeted during each production year to reduce liabilities and satisfy regulatory requirements prior to closure.

Waste dump cells are 100 x 100 x 15 m^3. The swell factor in this case study accounts for compaction within the waste dump facilities that is influenced by the equipment used for construction (Sari and Kumral 2018). An assumption made here is that the waste dump cell size allows for the model to adequately blend waste materials of different quality to prevent ARD. Considering the development of waste dumps forms in thin slices of rock, as trucks dump off the waste dump facility lifts, it is expected that adequate blending would occur with cell sizes of these dimensions. A NPR ratio, which measures the neutralization potential of mining wastes, is measured in each waste dump cell by accounting for the blended geochemical properties (Lawrence and Scheske 1997). NPR ratios greater then 2 are targeted in the waste dump cells to prevent the production of ARD. Geotechnical and operational constraints are considered when building the precedence relationships between waste-dump cells, which includes satisfying a 37.5-degree angle of repose and ensuring access is present from available roads. This is completed in a preprocessing step to define the precedence relationship in Eq. (5.20).

5.3.2 Simultaneous stochastic optimization with waste management and progressive reclamation

In the remainder of this section, the results from applying the proposed simultaneous stochastic optimization framework for long-term production are discussed and compared to a base case production schedule that does not consider integrated waste management and reclamation. The P-10, P-50, and P-90 represent a 10%, 50% and 90% chance of obtaining quantities below the values in the risk profiles that are presented throughout this section. The lower (LB) and upper (UB) bounds are drawn in the following figures when a constraint is applied. To begin with, the impacts of integrating waste management and progressive reclamation will be discussed to highlight the additional production scheduling considerations addressed in this case study. Then, the remaining production schedule decisions are discussed.

In Figures 5.3 and 5.4, the risk profiles for the NPR ratio at 51 randomly selected waste dump cells in the base case and integrated production schedule are shown. In the base case production schedule, there is significant risk of producing ARD as many waste dump cells have a blended NPR ratio that is less than the lower limit of 2. For example, the P-50 in waste dump cells 5-7 show that there is a large risk of producing acid in those areas of the waste dump due to the geochemical properties of waste placed. The integrated production schedule that considers waste management directly in the formulation is able to resolve these issues by blending materials to mitigate ARD. Although the method significantly reduces the risk of ARD and the frequency of occurrence, there are still deviations from the target NPR ratio, see cell 13 and 22 in Figure 5.4.

Alternative methods such as sourcing additional materials or placing limestone for neutralizing materials may also need to be considered to completely prevent ARD in the waste dump facilities. The other cells in the waste dump facilities show similar improvements when considering the integrated approach and reduce the number of cells with unsafe NPR ratios by 52.5%. This significantly reduces risk of ARD in the integrated production schedule with the proposed simultaneous stochastic optimization framework. The waste placement schedule is impacted by blending materials with different geochemical properties, which leads to distinct changes in the waste dump placement schedules in the base case versus the integrated production schedule shown in Figure 5.5.

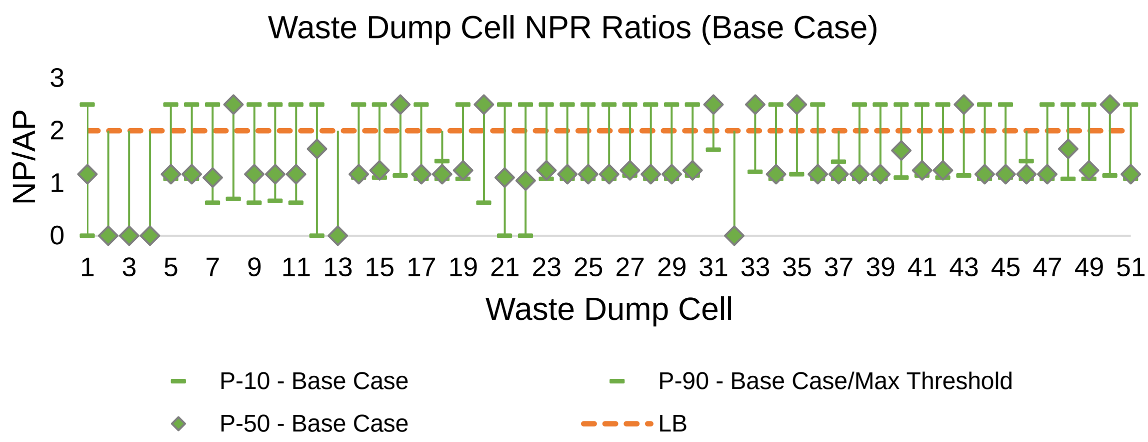


Figure 5.3. Risk profiles for cell NPR ratios in the waste dump cells in the base-case production schedule.

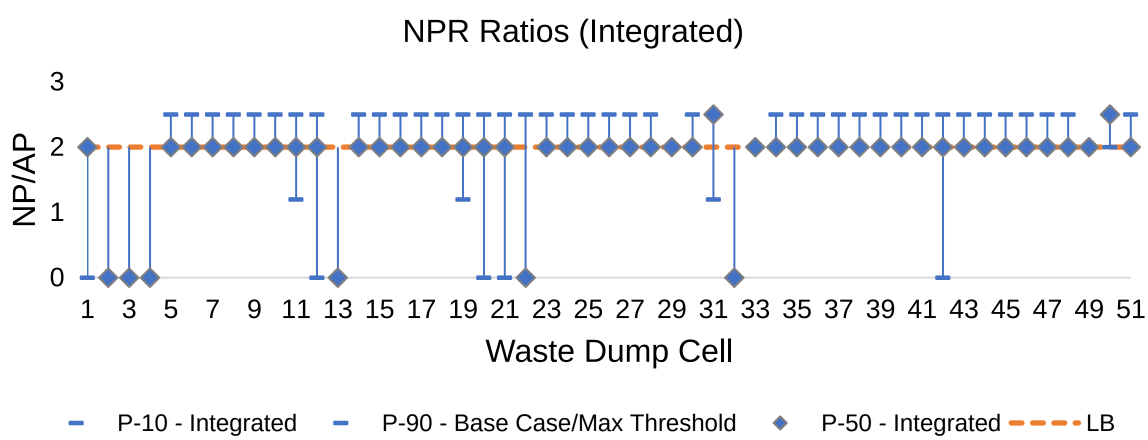


Figure 5.4. Risk profiles for cell NPR ratios in the waste dump cells in the integrated production schedule.

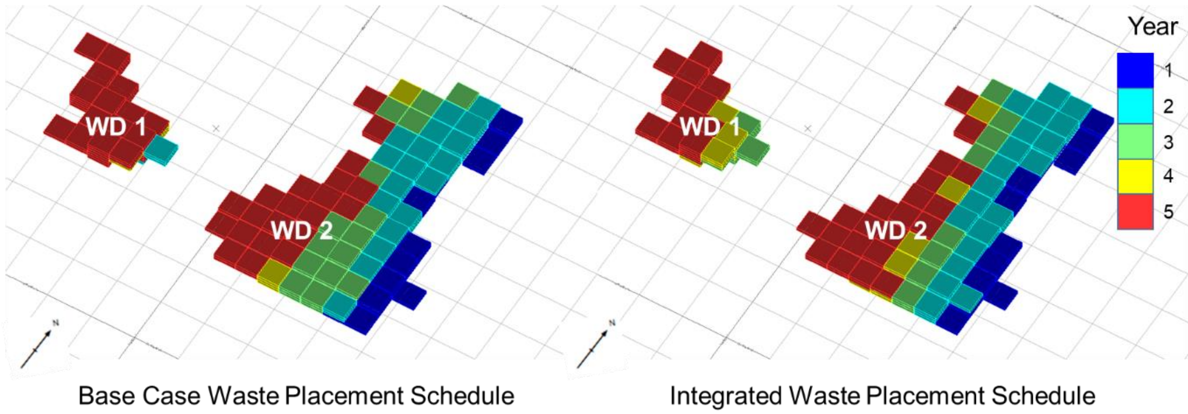


Figure 5.5. Base case (left) and integrated (right) waste dump placement schedule for waste dump facility (WD) 1 and 2.

Progressive reclamation and rehabilitation of waste dump facilities is desired to accelerate the closure of the mining complex and return these areas to a productive environmental state. Overburden/topsoil material covers were to be directly placed at the waste dump facilities to enhance plant growth and rehabilitation. However, the waste dump facility design must be considered, and full utilization of the available waste dump cells is required before starting the reclamation process to efficiently use the existing waste dump footprints. Figure 5.6 shows that the progressive reclamation process in the mining complex is unable to meet targets in years 1-2 in both production schedules. Progressive reclamation was ignored in previous long-term optimization studies leading to many areas of the waste dump being used and a lack of focus on completing areas to begin reclamation. The proposed optimization model targets areas of the waste dump to place waste material that provide opportunities to begin reclamation and satisfy the target 50 ha per a year. As the waste dump cells are filled to capacity, reclamation can begin in the areas that reach the final design and targets can be more easily meet in both production schedules in years 3-5. The advantage of integrating progressive reclamation into the optimization of mining complexes demonstrated in this application is that it provides engineering support for advancing the waste dump facilities to fulfill reclamation activities. In addition, Figure 5.7 shows that greater then 60% of reclamation activities on waste dumps 1 and 2 can be completed in parallel with production, if scheduled appropriately.

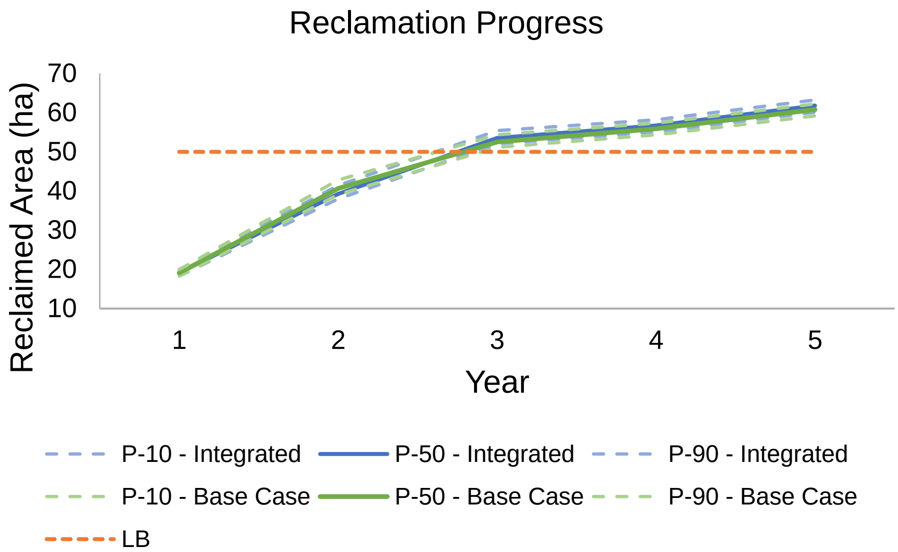


Figure 5.6. Risk profiles for the integrated (blue) and base-case (green) reclamation progress at the waste dump facilities.

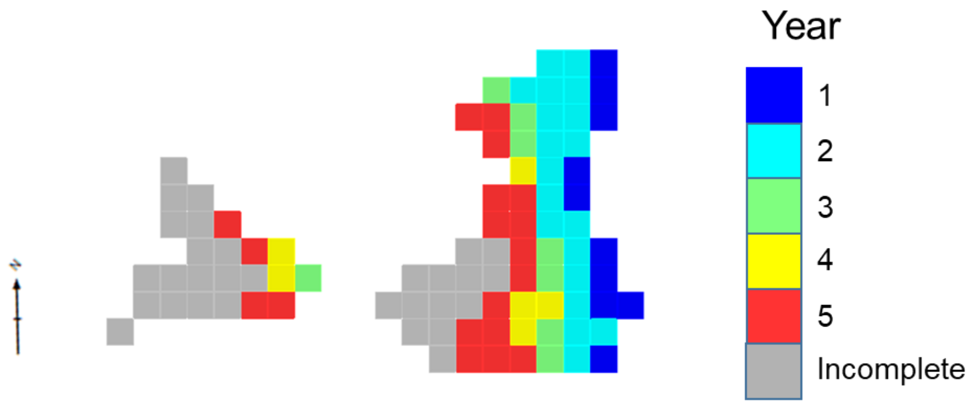


Figure 5.7. Progressive reclamation progress in the integrated production schedule (left) waste dump 1 and (right) waste dump 2.

As a result of introducing these additional production scheduling constraints into the stochastic programming formulation, there are impacts to the optimized production scheduling decisions. In Figure 5.8, the risk profiles for the processing facility throughput, mining production rate, and copper head grade, and copper production at the processing facility are shown. There are larger changes to optimized mining rate that occur in years 4 and 5. In addition, the processing facility throughput varies in each of the production schedules presented. A higher throughput with lower head grade is found at the processing plant with the integrated production schedule that considers waste management and reclamation. However, with the change in production scheduling decisions, the forecasted cash flow and metal production remain similar between the two production schedules. The cumulative discounted cash flow shown in Figure 5.9 demonstrates that the mining complex can break-even over the 5 production years while, preventing ARD and preparing for closure with additional reclamation costs incurred. Similar production cash flow and net present values are obtained; however, the integrated schedule allows for additional considerations that improve waste management and ensure progressive reclamation is completed.

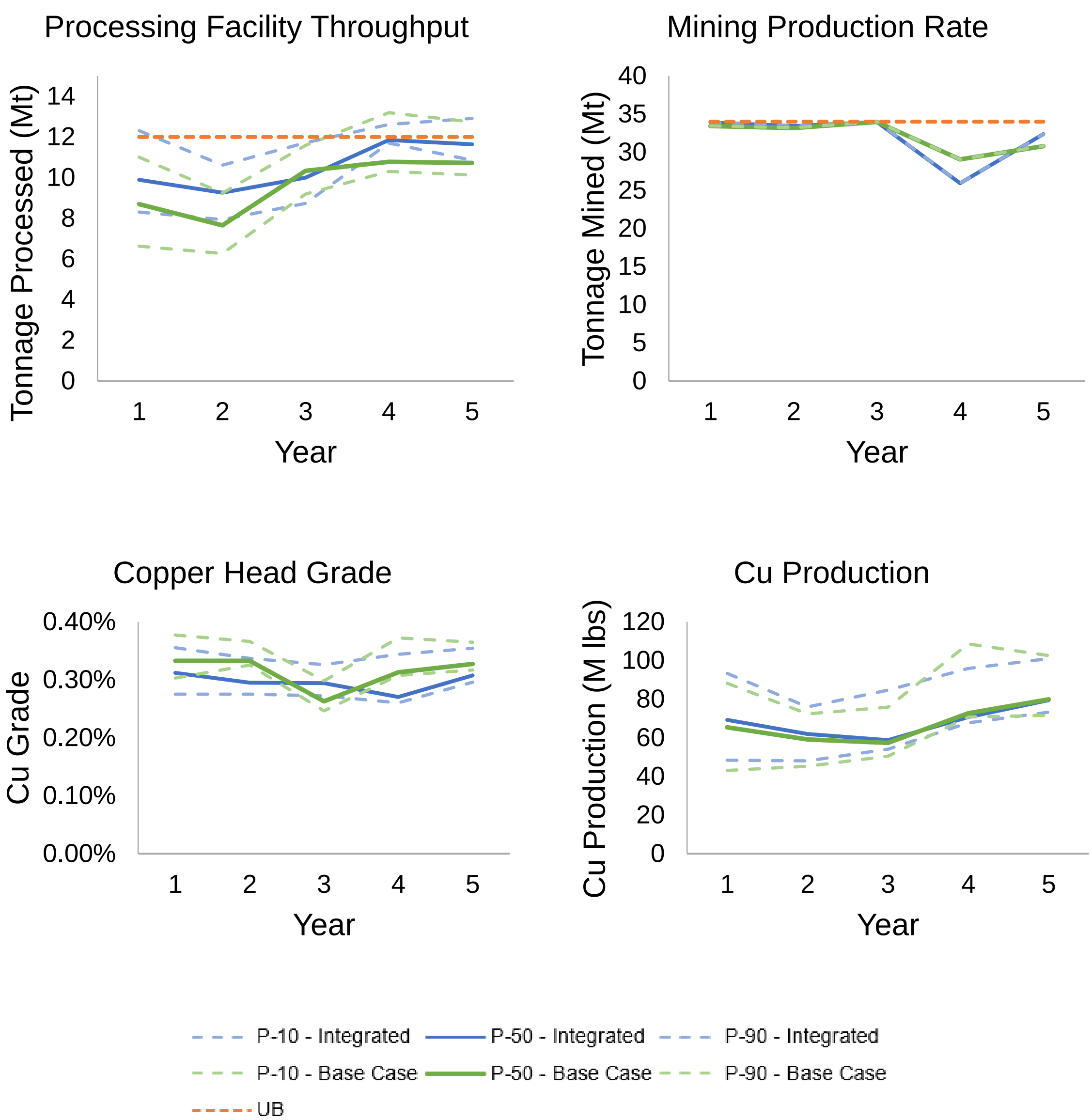


Figure 5.8. Risk profiles for the integrated and base-case for the processing plant annual throughput, mining production rate, copper head grade, and copper production.

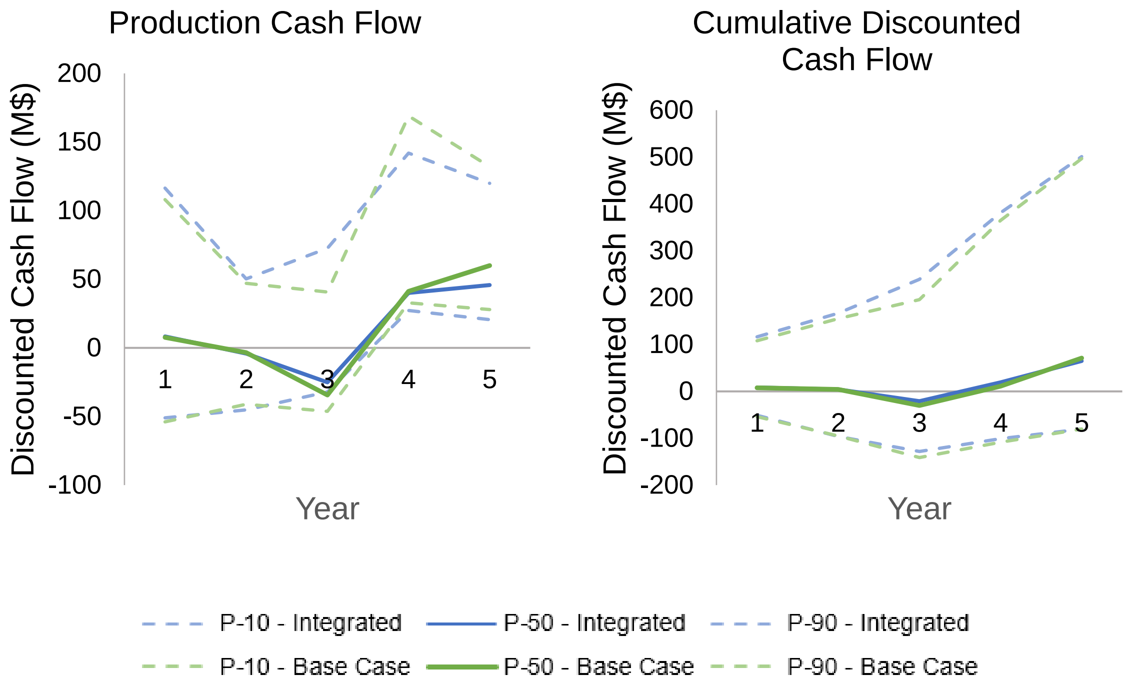


Figure 5.9. Risk profiles for the integrated and base-case discounted and cumulative discounted cash flow.

5.3.3 Discussion

The proposed optimization framework simultaneously optimizes several major components in a mining complex considering uncertainty. This leads to several critical differences in the resulting production schedules and also the mining extraction sequences in each mine. Cross sections for Mines 1-3 are shown in Figure 5.10 and 5.11, which highlight major differences between the extraction sequences produced when comparing the base case to the integrated schedule. The small arrows identify some of the visual differences in the schedule for Mine 1 and Mine 2. These changes in the extraction sequence impact the various subsequent components and allow for improved reclamation and waste management by maximizing the value of the products sold and accounting for the impact of the various critical attributes across the mining complex.

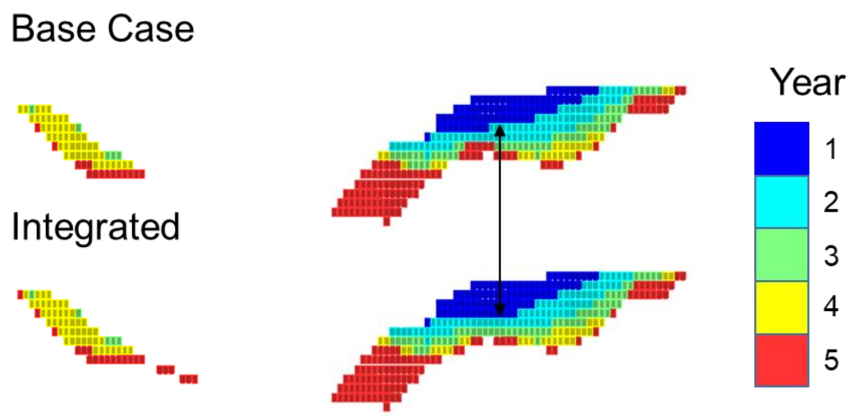


Figure 5.10. Comparison of the differences between the base case (top) and integrated (bottom) extraction sequence for Mine 1, an east-west cross-section.

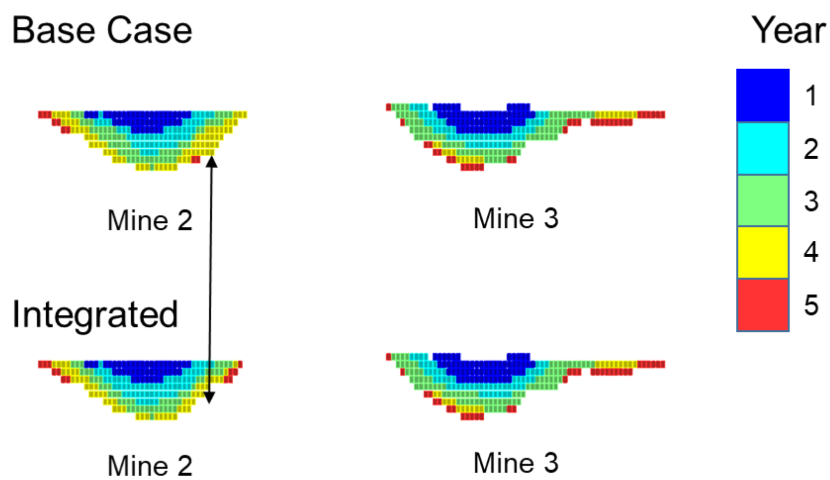


Figure 5.11. Comparison of the differences between the base case (top) and integrated (bottom) extraction sequence for Mine 2 and 3, an east-west cross section.

Optimizing long-term production schedules in industry scale mining complexes is a challenging optimization problem especially when considering several additional components for integrating waste management and progressive reclamation. Contextual bandits were explored in this work as a method for effectively applying metaheuristics to optimize mining complexes. In this case study, the results of optimizing with contextual bandits and an ϵ-greedy policy is compared to using a random approach for selecting perturbations within the simulated annealing framework. Figure

5.12 shows that a 24% higher objective function value is reached after 10,000 iterations and that the contextual bandit learnings are shown to improve the heuristic selection approach after only 2,000 iterations of the simulated annealing algorithm. Context related to the areas to be mined allows for improved efficiency in the optimization approach by understanding which areas are more likely to improve the production schedule by delaying them or advancing them in the extraction sequence.

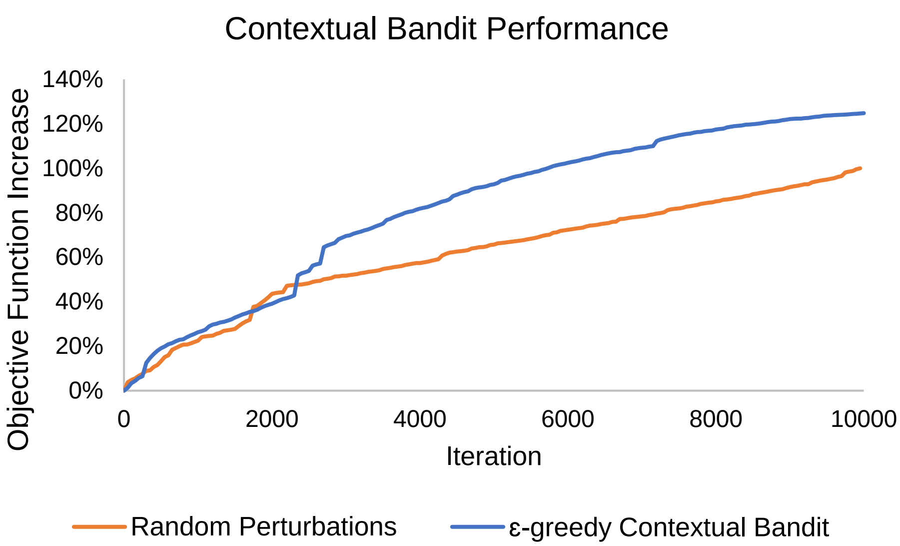


Figure 5.12. Contextual bandit performance compared to random perturbation selection.

The large scale simultaneous stochastic optimization framework contains over 9 M decision variables and 90 M constraints and can grow significantly depending on the number of mining blocks and waste dump cells considered within the optimization framework primarily due to the precedence relationships. This case study required 8 days to optimize to retrieve these results. Future work should consider further methods for improving the metaheuristic selection process to accelerate the optimization process.

5.4 Conclusions

A simultaneous stochastic optimization framework for optimizing large-scale mining complexes is proposed that integrates waste management and progressive reclamation considerations into the stochastic programming model. The framework directly manages supply uncertainty. A simulated annealing metaheuristic framework is applied to solve the optimization model with contextual bandits, which is tested at an open-pit copper-gold mining complex. The case study compares the simultaneous stochastic optimization with integrated waste management and progressive reclamation to a base case production schedule that does not consider these additional components in the optimization. The comparisons show that similar net present value can be achieved while, accounting for waste management and mitigating ARD by blending of materials with different geochemical properties. The integrated approach reduces the number of cells at risk of ARD by 52.5%. Progressive reclamation considerations allow for 60% of reclamation activities to be completed on waste dumps 1 and 2 during production. In addition, a reclamation and waste dump placement schedule are obtained that improve environmental performance. Lastly, contextual bandits are shown to increase the efficiency of the solution approach by selecting heuristics that are expected to improve the long-term production schedule based on context directly taken from the optimization of the case-study presented. Future work should investigate the integration of production cycle times particularly with regards to the load and haul fleet in open-pit mining complexes to accurately understand the impact of waste management and reclamation on fleet performance. In addition, further developments for improving metaheuristics through learning-based methods could help further improve the solution approach.

6 Conclusions and future work

Connecting planning horizons in mining complexes is critical for ensuring that long-term production targets can be obtained at shorter timescales without jeopardizing value. The work completed in this thesis has provided an integrative planning framework that combines short- and long-term planning timescales using stochastic programming and reinforcement learning. In addition, several critical production scheduling components have been integrated to improve operational and environmental performance in mining complexes for short- and long-term production scheduling. Preconcentration facilities provide additional benefits by eliminating waste from the process plant feed at a coarser scale then traditional processing methods, which improves process plant efficiency and provides further flexibility in short-term production schedules. Waste management techniques such as blending waste material with varying geochemical properties to mitigate acid rock drainage and create safe non-acid producing waste helps improve environmental performance. Lastly, progressive reclamation ensures that waste dump facilities are returned to a productive environmental state in parallel with production and allows for time to test and assess different reclamation methods. This can potentially reduce a major liability in operating mining complexes. These critical components are optimized simultaneously to capitalize on advantageous synergies between components when they exist.

The proposed optimization formulations and optimization techniques explored in this thesis are driven by the inputs; a set of stochastic orebody simulations of the materials in the ground. The stochastic orebody simulations are used to represent the mineral deposits in the mines and quantify the uncertainty and local variability of the critical material attributes. The information used to create the stochastic orebody simulation of the mineral deposits in a mining complex are obtained using geostatistical techniques that reproduce the statistics of the sampled exploration drillhole data. Infill drilling provides information that can be used to update the input stochastic orebody simulations and account for the data collected from different areas of the mineral deposits. An infill drilling optimization approach is developed in Chapter 2 to provide a strategic method to collect information and further understand the uncertainty and local variability of the material attributes in the mineral deposits, which directly impacts the related production schedules. The value of this additional information is assessed and infill drilling should be considered if further value can be obtained by adapting the long-term production scheduling decisions to take advantage of new information.

In Chapter 3, a short-term simultaneous stochastic optimization approach is developed that integrates preconcentration decisions. The approach aims to ensure that long-term production targets are met, while maximizing value and managing technical risk. A stochastic programming model and optimization approach are developed that allows for the simultaneous stochastic optimization of the extraction sequence, destination policy, stockpiling and preconcentration decisions. Actor-critic reinforcement learning and stochastic programming are applied to optimize the short-term production schedule and account for accessibility directly in the parametrization of the reinforcement learning agent's actions. A case study is completed in a copper mining complex with multiple mining areas, several material properties, stockpiles, preconcentration facilities, leach pads, process plants and waste dumps. The case study shows the practical aspects of the proposed optimization formulation and the integration of preconcentration, which led to a $140M increase in the first-year cash flow. Additionally, the actor–critic reinforcement learning algorithm learns a stable policy that is later used for assessing short-term production schedules in Chapter 4.

Chapter 4 presents a novel optimization framework for connecting short- and long-term production schedules in mining complexes. Planning across timescales is essential for optimizing mining complexes under uncertainty to eliminate risk of misalignment between long-term production forecasts and the outcome of the short-term production schedule. A stochastic mathematical programming framework is developed with an objective function that accounts for short- and long-term production and minimizes misalignment. The optimization framework leverages the embedded reinforcement learning approach developed in Chapter 3 and integrates it within a simulated annealing solution approach for rapidly assessing short-term production schedules using a learned neural network policy. The heuristic approach prevents the need to reoptimize the entire short-term schedule each time the long-term production is modified during optimization. The major advantage of this method is that the reinforcement learning approach continues to learn throughout the optimization, which allows the policy to adapt to new areas of the mineral deposit as the long-term production schedule is modified with the simulated annealing framework. An application is completed in a mining complex which shows that the approach reduces deviations from short-term operating cash flow by 21%. This is obtained by jointly optimizing the short- and long-term production schedules.

Chapter 5 extends the simultaneous stochastic optimization framework for long-term production scheduling in mining complexes by integrating waste management and progressive reclamation into the optimization framework. Waste management and progressive reclamation were considered by jointly optimizing the waste dump placement and reclamation schedule with the remaining production scheduling decisions. Blending materials with different geochemical properties is shown as a possible method for mitigating acid rock drainage in waste dump facilities. This is accomplished by accounting for material uncertainty and determining the appropriate blend of materials with varying qualities within the waste dump cells to prevent acid rock drainage. The simultaneous stochastic optimization approach is tested in a copper-gold mining complex and demonstrates significant improvements to environmental performance. Acid rock drainage is decreased by 52.5%, based on the number of waste dump cells that were originally acid generating given the blended NPR ratio. Progressive reclamation allowed for 60% of reclamation activities to be completed at two waste dumps facilities during production at a mining complex that was approaching closure, which helped reduce long-term liabilities. Lastly, contextual bandits were applied to improve the selection of perturbations within the optimization framework by learning the areas of each mine that should be advanced or delayed. The approach demonstrated a 24% higher objective function after only 10,000 iterations of the simulated annealing solution approach that was applied.

This thesis has demonstrated the continued importance of simultaneous stochastic optimization for mining complexes. By accounting for uncertainty and optimizing components together, significant value can be unlocked, and more realistic and operational production plans can be obtained. Reinforcement learning and stochastic programming were explored to optimize mining complexes. Stochastic programming allows for robust short- and long-term production schedules to be optimized under uncertainty, while directly managing risk in the optimization formulation. Reinforcement learning developments, including contextual bandits, show the benefits of learning a policy that can be used to improve the optimization formulation. Policies that account for key criteria related to the optimization formulation have been shown to improve the ability to: (1) select infill drilling locations in an uncertain environment; (2) select extraction sequence decisions based on the distribution of grades within the mining complex; and (3) learn the impact of advancing or delaying different areas of the mining complex based on related production scheduling information.

Future work related to the optimization of mining complexes is vast and requires continuous development. For appropriate infill drillhole planning, the proposed work should be extended to consider varying azimuth and dip directions for drilling. Additionally, extending the framework to consider the benefits of collecting information for a wider range of material attributes including geochemical and grade-by-size properties would be advantageous. This is because value is not solely driven by the valuable attributes of interest but also secondary properties that can influence screening performance and environmental components within the mining complex. Scheduling decisions based on further characterization of these attributes could lead to additional synergies that are often not considered during the optimization process and this requires further research.

Preconcentration via screening is only one method for separating materials based on the material grade-by-size properties. Future research should also investigate the integration of other sorting and separating methods that can reject waste at coarser scales then traditional processing facilities. These methods can take advantage of identifiable material properties that can be sensed using new and existing technologies and rejected prior to energy and water intensive processes. Appropriately quantifying the benefits of preconcentration requires two major steps. First, the characterization of the uncertain material properties that are expected to be leveraged by the preconcentration technique applied. Second, the integration of these components into the stochastic programming formulations for mining complexes and the definition of the appropriate decision variables that can be used to drive additional value. These integrations are complex and non-linear, as it is expected that several of these preconcentration techniques may need to be considered together to fully benefit from these technologies. This is a challenging and a very interesting area of future research for optimizing mining complexes as industry aims to improve environmental performance and increase efficiencies.

With regards to connecting planning timescales, the proposed method in this thesis is a static optimization formulation in the sense that a single optimization is completed to produce the short- and long-term production schedules. Updating frameworks that can use grade-control and production data to update the stochastic orebody simulations is critical for improving short- and long-term production schedules. Future work should consider an integrated approach where an updating framework is used to enhance the understanding of the mineral deposits and reoptimize the short- and long-term production plan on the fly. Online updating of production schedules with realignment of long-term production targets using an integrated optimization approach is expected to improve production guidance and provide better guidance. Additionally, the proposed method for connecting planning horizons was limited to only scheduling the first year using the embedded reinforcement learning agent. This was largely due to the computational overhead of optimizing these production schedules. Further research should investigate parallel computing opportunities for distributing the years of the long-term production schedule to several short-term production scheduling agents to accurately assess the value of the long-term production schedule when accessibility is considered. Different reinforcement learning algorithms would need to be explored to update the policy in this manner. Lastly, this method is applied considering production years and months but the effects across shorter timescales including weeks, days or hours could also be examined.

The optimization frameworks developed in this thesis for short- and long-term production scheduling do not consider cycle times particularly with regards to the load and haul fleet in open-pit mining complexes. For progressive reclamation and waste management to be successful it is important to understand the impact of fleet performance on production scheduling decisions. Ensuring that the load and haul fleet can meet the blending and reclamation requirements is essential to fulfill the plan. These considerations should be included in the optimization formulation to further inform the optimization process.

References

Aanonsen, S I, Nœvdal, G, Oliver, D S, Reynolds, A C, and Vallès, B (2009). The ensemble kalman filter in reservoir engineering—a review. *SPE Journal* 14(03): 393-412.

Adair, B, Revell, P, King, D, and Keeney, L (2020). A 'no tailings dam future' - preconcentration strategies 'down the sizes and across the value chain'. In: *Preconcentration Digital Conference 2020*, AusIMM Publication Series, Online.

Albor Consuegra, F and Dimitrakopoulos, R (2009). Stochastic mine design optimisation based on simulated annealing: Pit limits, production schedules, multiple orebody scenarios and sensitivity analysis. *Mining Technology* 118(2): 79-90.

Asif, Z and Chen, Z (2016). Environmental management in north american mining sector. *Environmental Science and Pollution Research* 23(1): 167-179.

Barbaro, R and Ramani, R (1986). Generalized multiperiod mip model for production scheduling and processing facilities selection and location. *Mining Engineering* 38(2): 107-114.

Ben-Awuah, E and Askari-Nasab, H (2013). Incorporating waste management into oil sands long term production planning. *Mining Technology* 122(1): 33-45.

Bengio, Y, Lodi, A, and Prouvost, A (2021). Machine learning for combinatorial optimization: A methodological tour d'horizon. *European Journal of Operational Research* 290(2): 405-421.

Benndorf, J (2014). Making use of online production data: Sequential updating of mineral resource models. *Mathematical Geosciences* 47(5): 547-563.

Benndorf, J (2020). Closed loop management in mineral resource extraction: Turning online geo-data into mining intelligence. Springer, Cham.

Benndorf, J and Buxton, M W N (2016). Sensor-based real-time resource model reconciliation for improved mine production control – a conceptual framework. *Mining Technology* 125(1): 54-64.

Benndorf, J and Dimitrakopoulos, R (2013). Stochastic long-term production scheduling of iron ore deposits: Integrating joint multi-element geological uncertainty. *Journal of Mining Science* 49(1): 68-81.

Birge, J R and Louveaux, F (2011). Introduction to stochastic programming. Springer, New York, NY.

Blom, M, Pearce, A R, and Stuckey, P J (2017). Short-term scheduling of an open-pit mine with multiple objectives. *Engineering Optimization* 49(5): 777-795.

Blom, M, Pearce, A R, and Stuckey, P J (2018). Multi-objective short-term production scheduling for open-pit mines: A hierarchical decomposition-based algorithm. *Engineering Optimization* 50(12): 2143-2160.

Blom, M, Pearce, A R, and Stuckey, P J (2019). Short-term planning for open pit mines: A review. *International Journal of Mining, Reclamation and Environment* 33(5): 318-339.

Blom, M L, Pearce, A R, and Stuckey, P J (2016). A decomposition-based algorithm for the scheduling of open-pit networks over multiple time periods. *Management Science* 62(10): 3059-3084.

Bolan, N, Kirkham, M B, and Ok, Y N-S (2017). Spoil to soil : Mine site rehabilitation and revegetation. Taylor & Francis, Boca Raton, FL.

Boland, N, Dumitrescu, I, and Froyland, G (2008). A multistage stochastic programming approach to open pit mine production scheduling with uncertain geology. *Optimization online*: 1-33.

Bolger, J (2000). Semi-quantitative laser-induced breakdown spectroscopy for analysis of mineral drill core. *Applied Spectroscopy* 54(2): 181-189.

Both, C and Dimitrakopoulos, R (2020). Joint stochastic short-term production scheduling and fleet management optimization for mining complexes. *Optimization and Engineering* 21(4): 1717-1743.

Both, C and Dimitrakopoulos, R (2021). Applied machine learning for geometallurgical throughput prediction—a case study using production data at the tropicana gold mining complex. *Minerals* 11(11): 1257.

Both, C and Dimitrakopoulos, R (2023a). Integrating geometallurgical ball mill throughput predictions into short-term stochastic production scheduling in mining complexes. *International Journal of Mining Science and Technology* 33(2): 185-199.

Both, C and Dimitrakopoulos, R (2023b). Utilisation of geometallurgical predictions of processing plant reagents and consumables for production scheduling under uncertainty. *International Journal of Mining, Reclamation and Environment* 37(1): 21-42.

Boucher, A and Dimitrakopoulos, R (2009). Block simulation of multiple correlated variables. *Mathematical Geosciences* 41(2): 215-237.

Boucher, A, Dimitrakopoulos, R, and Vargas-Guzmán, J (2005). Joint simulations, optimal drillhole spacing and the role of the stockpile. In: *Leuangthong, O and Deutsch, C V (eds). Geostatistics banff 2004*. Springer, Dordrecht: 35-44.

Bowman, D and Bearman, R (2014). Coarse waste rejection through size based separation. *Minerals Engineering* 62(7): 102-110.

Bratvold, R B, Bickel, J E, and Lohne, H P (2009). Value of information in the oil and gas industry: Past, present, and future. *SPE Reservoir Evaluation & Engineering* 12(04): 630-638.

Brika, Z, Gamache, M, and Dimitrakopoulos, R (2023). Optimising the mine production scheduling accounting for stockpiling and investment decisions under geological uncertainty. *International Journal of Mining, Reclamation and Environment*: 1-20.

Burns, R and Grimes, A (1986). The application of preconcentration by screening at b.C.L. In: *AusIMM Mineral Development Symposium*, AusIMM, Madang, PNG.

Caccetta, L (2007). Application of optimisation techniques in open pit mining. In: *Weintraub, A, Romero, C, Bjørndal, T, Epstein, R and Miranda, J (eds). Handbook of operations research in natural resources*. Springer US, Boston, MA: 547-559.

Carrasco, C, Keeney, L, and Napier-Munn, T (2016a). Methodology to develop a coarse liberation model based on preferential grade by size responses. *Minerals Engineering* 86(2): 149-155.

Carrasco, C, Keeney, L, Napier-Munn, T J, and Bode, P (2017). Unlocking additional value by optimising comminution strategies to process grade engineering® streams. *Minerals Engineering* 103-104(4): 2-10.

Carrasco, C, Keeney, L, and Walters, S (2016b). Development of a novel methodology to characterise preferential grade by size deportment and its operational significance. *Minerals Engineering* 91(5): 100-107.

Chanda, E C K (1990). An application of goal programming to production planning in the crushed stone industry. *International Journal of Surface Mining, Reclamation and Environment* 4(3): 125-129.

Chanda, E K C and Dagdelen, K (1995). Optimal blending of mine production using goal programming and interactive graphics systems. *International Journal of Surface Mining, Reclamation and Environment* 9(4): 203-208.

Chen, Y, Oliver, D S, and Zhang, D (2009). Data assimilation for nonlinear problems by ensemble kalman filter with reparameterization. *Journal of Petroleum Science and Engineering* 66(1): 1-14.

Cressie, N (1991). Statistics for spatial data. Wiley, New York.

Dagdelen, K (2001). Open pit optimization-strategies for improving economics of mining projects through mine planning. In: *17th International Mining Congress and Exhibition of Turkey*, Ankara, Turkey.

Day, S J (2022). Small-scale field evaluation of geochemical blending of waste rock to mitigate acid rock drainage potential. *Geochemistry: Exploration, Environment, Analysis* 22(1).

De Carvalho, J P and Dimitrakopoulos, R (2019). Effects of high-order simulations on the simultaneous stochastic optimization of mining complexes. *Minerals* 9(4): 210.

Del Castillo, M F (2018). Dynamic simultaneous optimization of mineral value chains under resource uncertainty. *PhD Thesis*, McGill University, Montreal, Quebec, Canada.

Del Castillo, M F and Dimitrakopoulos, R (2019). Dynamically optimizing the strategic plan of mining complexes under supply uncertainty. *Resources Policy* 60(3): 83-93.

Delalleau, O, Peter, M, Alonso, E, and Logut, A (2019). Discrete and continuous action representation for practical rl in video games. *arXiv preprint arXiv:1912.11077*.

Delmelle, E M and Goovaerts, P (2009). Second-phase sampling designs for non-stationary spatial variables. *Geoderma* 153(1-2): 205-216.

Dengiz, B, Altiparmak, F, and Smith, A E (1997). Local search genetic algorithm for optimal design of reliable networks. *IEEE transactions on Evolutionary Computation* 1(3): 179-188.

Diehl, P and David, M (1982). Classification of ore reserves/resources based on geostatistical methods. *CIM Bulletin* 75(838).

Dimitrakopoulos, R, Farrelly, C, and Godoy, M (2002). Moving forward from traditional optimization: Grade uncertainty and risk effects in open-pit design. *Mining Technology* 111(1): 82-88.

Dimitrakopoulos, R and Jewbali, A (2013). Joint stochastic optimisation of short and long term mine production planning: Method and application in a large operating gold mine. *Mining Technology* 122(2): 110-123.

Dimitrakopoulos, R and Lamghari, A (2022). Simultaneous stochastic optimization of mining complexes-mineral value chains: An overview of concepts, examples and comparisons. *International Journal of Mining, Reclamation and Environment* 36(6): 1-18.

Dimitrakopoulos, R, Mustapha, H, and 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(1): 65-99.

Dimitrakopoulos, R and Ramazan, S (2004). Uncertainty based production scheduling in open pit mining. *SME transactions* 316.

Dimitrakopoulos, R and Ramazan, S (2008). Stochastic integer programming for optimising long term production schedules of open pit mines: Methods, application and value of stochastic solutions. *Mining Technology* 117(4): 155-160.

Dirkx, R and Dimitrakopoulos, R (2018). Optimizing infill drilling decisions using multi-armed bandits: Application in a long-term, multi-element stockpile. *Mathematical Geosciences* 50(1): 35-52.

Dowd, P (1994). Risk assessment in reserve estimation and open-pit planning. *Transactions of the Institution of Mining and Metallurgy Section A: Mining Industry* 103: A148-154.

Dowd, P (1997). Risk in minerals projects: Analysis, perception and management. In: *Transactions of the Institution of Mining and Metallurgy Section A: Mining Industry*.

Eivazy, H and Askari-Nasab, H (2012). A mixed integer linear programming model for short-term open pit mine production scheduling. *Mining Technology* 121(2): 97-108.

Espejel, C, Scott, M, and Nehring, M (2017). Economic benefits and technical complexities of grade engineering® in strategic mine planning of metalliferous projects. In: *Proceedings of the International Conference Mine Planning and Equipment Selection*, Dresden, Germany.

Evensen, G (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research* 99(C5).

Evensen, G (2003). The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics* 53(4): 343-367.

Fathollahzadeh, K, Mardaneh, E, Cigla, M, and Asad, M W A (2021). A mathematical model for open pit mine production scheduling with grade engineering® and stockpiling. *International Journal of Mining Science and Technology* 31(4): 717-728.

Ferland, J A, Amaya, J, and Djuimo, M S (2007). Application of a particle swarm algorithm to the capacitated open pit mining problem. In: *Mukhopadhyay, S C and Gupta, G S (eds). Autonomous robots and agents*. Springer, Berlin, Heidelberg: 127-133.

Froyland, G, Menabde, M, Stone, P, and Hodson, D (2018). The value of additional drilling to open pit mining projects. In: *Dimitrakopoulos, R (eds). Advances in applied strategic mine planning*. Springer, Cham: 119-138.

Fu, Z, Asad, M W A, and Topal, E (2019). A new model for open-pit production and waste-dump scheduling. *Engineering Optimization* 51(4): 718-732.

Fu, Z, Li, Y, Topal, E, and Williams, D (2015). A new tool for optimisation of mine waste management in potential acid forming conditions. In: *Tailings and Mine Waste Management for the 21st Century*, Sydney, Australia.

Fytas, K, Hadjigeorgiou, J, and Collins, J (1993). Production scheduling optimization in open pit mines. *International Journal of Surface Mining and Reclamation* 7(1): 1-9.

Fytas, K, Pelley, C, and Calder, P (1987). Optimization of open pit short- and long-range production scheduling. *CIM Bulletin* 80(904): 55-61.

Gershon, M (1983). Optimal mine production scheduling: Evaluation of large scale mathematical programming approaches. *International Journal of Mining Engineering* 1(4): 315-329.

Gershon, M, Allen, L E, and Manley, G (1988). Application of a new approach for drillholes location optimization. *International Journal of Surface Mining, Reclamation and Environment* 2(1): 27-31.

Godoy, M (2003). The effective management of geological risk in long-term scheduling of open pit mines. *Ph.D. Thesis*, University of Queensland, Brisbane, Australia.

Goldberg, D E (2013). Genetic algorithms. Pearson Education, India.

Goodfellow, R and Dimitrakopoulos, R (2015). Stochastic optimization of open-pit mining complexes with capital expenditures: Application at a copper mining complex. In: *Applications of Computer and Operations Research in the Mineral Industry (APCOM)*, Fairbanks, AK.

Goodfellow, R and Dimitrakopoulos, R (2016). Global optimization of open pit mining complexes with uncertainty. *Applied Soft Computing* 40(3): 292-304.

Goodfellow, R and Dimitrakopoulos, R (2017). Simultaneous stochastic optimization of mining complexes and mineral value chains. *Mathematical Geosciences* 49(3): 341-360.

Goovaerts, P (1997). Geostatistics for natural resources evaluation. Oxford University Press, New York.

Goria, S, Armstrong, M, and Galli, A (2001). Quanitfying the impact of additional drilling on an open-pit gold project. In: *IAMG2001*, Cancun, Mexico.

Grant, C and Koch, J (2007). Decommissioning western australia's first bauxite mine: Co-evolving vegetation restoration techniques and targets. *Ecological Management & Restoration* 8(2): 92-105.

Grimalt, J O, Ferrer, M, and Macpherson, E (1999). The mine tailing accident in aznalcollar. *Science of the Total Environment* 242(1-3): 3-11.

Haarnoja, T, Zhou, A, Abbeel, P, and Levine, S (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: *International conference on machine learning*, PMLR, Stockholm, Sweden.

Hilscher, B, Nayak, P, Lorio, L, and Yoon, N (2017). Technical review and evaluation of ore sorting technologies on precious metals operations. *49th COM, Vancouver, BC, Canada*.

Hoerger, S, Bachmann, J, Criss, K, and Shortridge, E (1999a). Long term mine and process scheduling at newmont's nevada operations. In: *Proceedings of the 28th Symposium on the Application of Computers and Operations Research in the Mineral Industry*.

Hoerger, S, Seymour, F, and Hoffman, L (1999b). Mine planning at newmont's nevada operations. *Mining engineering* 51(10): 26-30.

Holmes, P M (2001). Shrubland restoration following woody alien invasion and mining: Effects of topsoil depth, seed source, and fertilizer addition. *Restoration Ecology* 9(1): 71-84.

Huo, D, Sari, Y A, Kealey, R, and Zhang, Q (2023). Reinforcement learning-based fleet dispatching for greenhouse gas emission reduction in open-pit mining operations. *Resources, Conservation and Recycling* 188(1): 106664.

Hustrulid, W, Kuchta, M, and Martin, R (2013). Open pit mine planning & design. CRC Press, Boca Raton, FL.

Iandola, F N, Han, S, Moskewicz, M W, Ashraf, K, Dally, W J, and Keutzer, K (2016). Squeezenet: Alexnet-level accuracy with 50x fewer parameters and< 0.5 mb model size. *arXiv preprint arXiv:1602.07360*.

Ibm (2017). Ibm ilog cplex optimization studio cplex user's manual.

Isaaks, E H (1991). The application of monte carlo methods to the analysis of spatially correlated data. *Ph.D. Thesis*, Stanford University, Stanford, CA.

Isaaks, E H and Srivastava, R M (1989). Applied geostatistics. Oxford university press New York.

Jazwinski, A H (1970). Stochastic processes and filtering theory. Academic Press, New York.

Jewbali, A and Dimitrakopoulos, R (2011). Implementation of conditional simulation by successive residuals. *Computers & Geosciences* 37(2): 129-142.

Journal, A G and Alabert, F (1989). Non-gaussian data expansion in the earth sciences. *Terra Nova* 1(2): 123-134.

Journel, A (2002). Roadblocks to the evaluation of ore reserves — the simulation overpass and putting more geology into numerical models of deposits. In: *Dimitrakopoulos, R (eds). Orebody modelling and strategic mine planning*. Specturm Series. 14: 29-32.

Kalman, R E (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82(1): 35-45.

Kennedy, J and Eberhart, R (1995). Particle swarm optimization. In: *Proceedings of ICNN'95 - International Conference on Neural Networks*.

Kingma, D P and Ba, J (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kirkpatrick, S, Gelatt Jr, C D, and Vecchi, M P (1983). Optimization by simulated annealing. *science* 220(4598): 671-680.

Kizilkale, A C and Dimitrakopoulos, R (2014). Optimizing mining rates under financial uncertainty in global mining complexes. *International Journal of Production Economics* 158: 359-365.

Koch, J M (2007). Restoring a jarrah forest understorey vegetation after bauxite mining in western australia. *Restoration Ecology* 15: S26-S39.

Kossoff, D, Dubbin, W, Alfredsson, M, Edwards, S, Macklin, M, and Hudson-Edwards, K A (2014). Mine tailings dams: Characteristics, failure, environmental impacts, and remediation. *Applied Geochemistry* 51: 229-245.

Kozan, E and Liu, S Q (2016). A new open-pit multi-stage mine production timetabling model for drilling, blasting and excavating operations. *Mining Technology* 125(1): 47-53.

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

Kumar, A and Dimitrakopoulos, R (2021). Production scheduling in industrial mining complexes with incoming new information using tree search and deep reinforcement learning. *Applied Soft Computing* 110(10): 107644.

Kumar, A, Dimitrakopoulos, R, and Maulen, M (2020). Adaptive self-learning mechanisms for updating short-term production decisions in an industrial mining complex. *Journal of Intelligent Manufacturing* 31(7): 1795-1811.

Kumar, D and Srinivasan, S (2019). Ensemble-based assimilation of nonlinearly related dynamic data in reservoir models exhibiting non-gaussian characteristics. *Mathematical Geosciences* 51(1): 75-107.

Kumar, D and Srinivasan, S (2020). Indicator-based data assimilation with multiple-point statistics for updating an ensemble of models with non-gaussian parameter distributions. *Advances in Water Resources* 141(10).

Kumral, M (2013). Optimizing ore–waste discrimination and block sequencing through simulated annealing. *Applied Soft Computing* 13(8): 3737-3744.

Kumral, M and Dowd, P (2002). Short-term mine production scheduling for industrial minerals using multi-objective simulated annealing. 2002. In: *Proceedings of 30th APCOM symp.: Fairbanks, Alaska*.

Kumral, M and Dowd, P A (2005). A simulated annealing approach to mine production scheduling. *Journal of the Operational Research Society* 56(8): 922-930.

Kumral, M and Ozer, U (2013). Planning additional drilling campaign using two-space genetic algorithm: A game theoretical approach. *Computers & Geosciences* 52(3): 117-125.

L'heureux, G, Gamache, M, and Soumis, F (2013). Mixed integer programming model for short term planning in open-pit mines. *Mining Technology* 122(2): 101-109.

La Rosa, D (2017). The development of a new geometrical blast fragmentation model and its application to grade engineering®. In: *European Federation of Explosives Engineers*, Stockholm.

Lamghari, A and 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(3): 642-652.

Lamghari, A and Dimitrakopoulos, R (2020). Hyper-heuristic approaches for strategic mine planning under uncertainty. *Computers & Operations Research* 115: 104590.

Lamghari, A, Dimitrakopoulos, R, and Ferland, J A (2014). A variable neighbourhood descent algorithm for the open-pit mine production scheduling problem with metal uncertainty. *Journal of the Operational Research Society* 65(9): 1305-1314.

Lamghari, A, Dimitrakopoulos, R, and Senécal, R (2022). A matheuristic approach for optimizing mineral value chains under uncertainty. *Optimization and Engineering* 23(2): 1139-1164.

Lane, K (1964). Choosing the optimum cut-foff grade. *(eds). Colorado school of mines quarterly report*. 59: 811-829.

Lane, K, Hamilton, D, and Parker, J (1984). Cutoff grades for two minerals. Symposium on the Application of Computers and Operations Research in the Minerals Industries**:** 485-490.

Laroche-Boisvert, M, Dimitrakopoulos, R, and Ferland, J A (2021). Simultaneous production scheduling and transportation optimization from mines to port under uncertain material supply. *Resources Policy* 73: 102150.

Lawrence, R W and Scheske, M (1997). A method to calculate the neutralization potential of mining wastes. *Environmental Geology* 32(2): 100-106.

Leite, A and Dimitrakopoulos, R (2007). Stochastic optimisation model for open pit mine planning: Application and risk analysis at copper deposit. *Mining Technology* 116(3): 109-118.

Leite, A and Dimitrakopoulos, R (2014). Stochastic optimization of mine production scheduling with uncertain ore/metal/waste supply. *International Journal of Mining Science and Technology* 24(6): 755-762.

Levinson, Z and Dimitrakopoulos, R (2020a). Adaptive simultaneous stochastic optimization of a gold mining complex: A case study. *Journal of the Southern African institute of mining and metallurgy* 120(3): 221-232.

Levinson, Z and Dimitrakopoulos, R (2020b). Simultaneous stochastic optimisation of an open-pit gold mining complex with waste management. *International Journal of Mining, Reclamation and Environment* 34(6): 415-429.

Levinson, Z, Dimitrakopoulos, R, and Keutchayan, J (2023). Simultaneous stochastic optimization of an open-pit mining complex with preconcentration using reinforcement learning. *Applied Soft Computing* 138(5): 110180.

Li, Y, Topal, E, and Williams, D (2013). Waste rock dumping optimisation using mixed integer programming (mip). *International Journal of Mining, Reclamation and Environment* 27(6): 425-436.

Li, Y, Topal, E, and Williams, D J (2014). Optimisation of waste rock placement using mixed integer programming. *Mining Technology* 123(4): 220-229.

Lillicrap, T P, Hunt, J J, Pritzel, A, Heess, N, Erez, T, Tassa, Y, Silver, D, and Wierstra, D (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*.

Lin, L-J (1992). Reinforcement learning for robots using neural networks. *Ph.D. Thesis*, Carnegie Mellon University, Pittsburgh, PA.

Liu, S Q and Kozan, E (2019). Integration of mathematical models for ore mining industry. *International Journal of Systems Science: Operations & Logistics* 6(1): 55-68.

Maas, A L, Hannun, A Y, and Ng, A Y (2013). Rectifier nonlinearities improve neural network acoustic models. In: *Proc. icml*, Atlanta, Georgia, USA.

Macdonald, S E, Landhäusser, S M, Skousen, J, Franklin, J, Frouz, J, Hall, S, Jacobs, D F, and Quideau, S (2015). Forest restoration following surface mining disturbance: Challenges and solutions. *New Forests* 46(5): 703-732.

Mackenzie, D D and Naeth, M A (2010). The role of the forest soil propagule bank in assisted natural recovery after oil sands mining. *Restoration Ecology* 18(4): 418-427.

Mai, N L, Topal, E, Erten, O, and Sommerville, B (2019). A new risk-based optimisation method for the iron ore production scheduling using stochastic integer programming. *Resources Policy* 62: 571-579.

Mai, N L, Topal, E, and Ertent, O (2018). A new open-pit mine planning optimization method using block aggregation and integer programming. *Journal of the Southern African Institute of Mining and Metallurgy* 118: 705-714.

Mann, C and Wilke, F L (1992). Open pit short-term mine planning for grade control - a combination of cad techniques and linear programming. In: *Applications of Computers and Operations Research in the Mineral Industyr - Proceedings of the 23rd International Symposium*.

Mariethoz, G and Caers, J (2014). Multiple-point geostatistics : Stochastic modeling with training images. Hoboken, Wiley.

Matamoros, M E V and Dimitrakopoulos, R (2016). Stochastic short-term mine production schedule accounting for fleet allocation, operational considerations and blending restrictions. *European Journal of Operational Research* 255(3): 911-921.

Mchaina, D M (2001). Environmental planning considerations for the decommissioning, closure and reclamation of a mine site. *International Journal of Surface Mining, Reclamation and Environment* 15(3): 163-176.

Mehling, P, Day, S, and Sexsmith, K (1997). Blending and layering waste rock to delay, mitigate or prevent acid generation. In: *Fourth International Conference on Acid Rock Drainage*, Vancouver, BC, Canada.

Menabde, M, Froyland, G, Stone, P, and Yeates, G (2007). Orebody modelling and strategic mine planning. The Australasian Institute of Mining and Metallurgy Melbourne, Australia.

Metropolis, N, Rosenbluth, A W, Rosenbluth, M N, Teller, A H, and Teller, E (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics* 21(6): 1087-1092.

Minniakhmetov, I and Dimitrakopoulos, R (2017a). A high-order, data-driven framework for joint simulation of categorical variables. In: *Gómez-Hernández, J J, Rodrigo-Ilarri, J, Rodrigo-Clavero, M E, Cassiraga, E and Vargas-Guzmán, J A (eds). Geostatistics valencia 2016*. Springer International Publishing, Cham: 287-301.

Minniakhmetov, I and Dimitrakopoulos, R (2017b). Joint high-order simulation of spatially correlated variables using high-order spatial statistics. *Mathematical Geosciences* 49(1): 39-66.

Minniakhmetov, I, Dimitrakopoulos, R, and Godoy, M (2018). High-order spatial simulation using legendre-like orthogonal splines. *Mathematical Geosciences* 50(7): 753-780.

Mitchell, M (1998). An introduction to genetic algorithms. MIT press.

Mnih, V, Badia, A P, Mirza, M, Graves, A, Lillicrap, T, Harley, T, Silver, D, and Kavukcuoglu, K (2016). Asynchronous methods for deep reinforcement learning. In: *International conference on machine learning*, PMLR.

Mnih, V, Kavukcuoglu, K, Silver, D, Graves, A, Antonoglou, I, Wierstra, D, and Riedmiller, M (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Mnih, V, Kavukcuoglu, K, Silver, D, Rusu, A A, Veness, J, Bellemare, M G, Graves, A, Riedmiller, M, Fidjeland, A K, and Ostrovski, G (2015). Human-level control through deep reinforcement learning. *nature* 518(7540): 529-533.

Montiel, L and Dimitrakopoulos, R (2015). Optimizing mining complexes with multiple processing and transportation alternatives: An uncertainty-based approach. *European Journal of Operational Research* 247(1): 166-178.

Montiel, L and Dimitrakopoulos, R (2017). A heuristic approach for the stochastic optimization of mine production schedules. *Journal of Heuristics* 23(5): 397-415.

Montiel, L and Dimitrakopoulos, R (2018). Simultaneous stochastic optimization of production scheduling at twin creeks mining complex, nevada. *Mining Engineering* 70(12): 48-56.

Morales, D and Dimitrakopoulos, R (2021). High-order simulation of geological units at the sara-macca gold deposit, suriname. *Les Cahiers du GERAD ISSN* 711: 2440.

Mousavi, A, Kozan, E, and Liu, S Q (2016a). Comparative analysis of three metaheuristics for short-term open pit block sequencing. *Journal of Heuristics* 22(3): 301-329.

Mousavi, A, Kozan, E, and Liu, S Q (2016b). Open-pit block sequencing optimization: A mathematical model and solution technique. *Engineering Optimization* 48(11): 1932-1950.

Mudd, G M (2007). Global trends in gold mining: Towards quantifying environmental and resource sustainability. *Resources Policy* 32(1): 42-56.

Mudd, G M (2012). Key trends in the resource sustainability of platinum group elements. *Ore Geology Reviews* 46(08): 106-117.

Mustapha, H and Dimitrakopoulos, R (2010). High-order stochastic simulation of complex spatially distributed natural phenomena. *Mathematical Geosciences* 42(5): 457-485.

Nadolski, S, Samuels, M, Klein, B, and Hart, C J R (2018). Evaluation of bulk and particle sensor-based sorting systems for the new afton block caving operation. *Minerals Engineering* 121(6): 169-179.

Nehring, M, Topal, E, Kizil, M, and Knights, P (2010). An investigation to integrate optimum long-term planning with short planning in underground mine production scheduling. In: *Mine Planning and Equipment Selection Conference*.

Neunert, M, Abdolmaleki, A, Wulfmeier, M, Lampe, T, Springenberg, T, Hafner, R, Romano, F, Buchli, J, Heess, N, and Riedmiller, M (2020). Continuous-discrete reinforcement learning for hybrid control in robotics. In: *Conference on Robot Learning*, PMLR.

Nordstrom, D K (2011). Mine waters: Acidic to circumneutral. *Elements* 7(6): 393-398.

Northey, S, Mohr, S, Mudd, G M, Weng, Z, and Giurco, D (2014). Modelling future copper ore grade decline based on a detailed assessment of copper resources and mining. *Resources, Conservation and Recycling* 83(2): 190-201.

Paduraru, C and Dimitrakopoulos, R (2018). Adaptive policies for short-term material flow optimization in a mining complex. *Mining Technology* 127(1): 56-63.

Paduraru, C and Dimitrakopoulos, R (2019). Responding to new information in a mining complex: Fast mechanisms using machine learning. *Mining Technology* 128(3): 129-142.

Paithankar, A, Chatterjee, S, Goodfellow, R, and Asad, M W A (2020). Simultaneous stochastic optimization of production sequence and dynamic cut-off grades in an open pit mining operation. *Resources Policy* 66: 101634.

Pan, S J and Yang, Q (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22(10): 1345-1359.

Parker, H M (2012). Reconciliation principles for the mining industry. *Mining Technology* 121(3): 160-176.

Parrot, A, Vanier, F, and Blouin, A (2020). Standoff mid-infrared reflectance spectroscopy using quantum cascade laser for mineral identification. In: *SPIE Future Sensing Technologies*, SPIE, Online.

Paszke, A, Gross, S, Massa, F, Lerer, A, Bradbury, J, Chanan, G, Killeen, T, Lin, Z, Gimelshein, N, and Antiga, L (2019). Pytorch: An imperative style, high-performance deep learning library. In: *Advances in neural information processing systems 2019*.

Peattie, R and Dimitrakopoulos, R (2013). Forecasting recoverable ore reserves and their uncertainty at morila gold deposit, mali: An efficient simulation approach and future grade control drilling. *Mathematical Geosciences* 45(8): 1005-1020.

Peters, J, Vijayakumar, S, and Schaal, S (2005). Natural actor-critic. In: *Gama, J, Camacho, R, Brazdil, P B, Jorge, A M and Torgo, L (eds). Machine learning: Ecml 2005*. Springer, Berlin, Heidelberg. 3720: 280-291.

Pimentel, B S, Mateus, G R, and Almeida, F A (2010). Mathematical models for optimizing the global mining supply chain. In: *Nag, B (eds). Intelligent systems in operations: Methods, models and applications in the supply chain*. IGI Global: 133-163.

Powell, W B (2022). Reinforcement learning and stochastic optimization: A unified framework for sequential decisions. John Wiley & Sons, Inc, Hoboken, NJ.

Price, W A (2003). Challenges posed by metal leaching and acid rock drainage at closed mines. In: *British Columbia Mine Reclamation Symposium*.

Quigley, M and Dimitrakopoulos, R (2020). Incorporating geological and equipment performance uncertainty while optimising short-term mine production schedules. *International Journal of Mining, Reclamation and Environment* 34(5): 362-383.

Ramazan, S and Dimitrakopoulos, R (2004). Stochastic optimisation of long-term production scheduling for open pit mines with a new integer programming formulation. In: *Dimitrakopoulos, R (eds). Orebody modelling and strategic mine planning*. Australasian Institute of Mining and Metallurgy (AusIMM), Carlton, Vic. 14.

Ramazan, S and Dimitrakopoulos, R (2013). Production scheduling with uncertain supply: A new solution to the open pit mining problem. *Optimization and engineering* 14(2): 361-380.

Ravenscroft, P (1992). Risk analysis for mine scheduling by conditional simulation. *Transactions of the Institution of Mining and Metallurgy. Section A. Mining Industry* 101: 104-108.

Remy, N, Boucher, A, and Wu, J (2009). Applied geostatistics with sgems: A user's guide. Cambridge University Press, Cambridge, Massachusetts.

Rimélé, A, Dimitrakopoulos, R, and Gamache, M (2018). A stochastic optimization method with in-pit waste and tailings disposal for open pit life-of-mine production planning. *Resources Policy* 57(8): 112-121.

Rossi, M E and Deutsch, C V (2014). Mineral resource estimation. Springer, Dordrecht.

Saliba, Z and Dimitrakopoulos, R (2019). Simultaneous stochastic optimization of an open pit gold mining complex with supply and market uncertainty. *Mining Technology* 128(4): 216-229.

Saliba, Z and Dimitrakopoulos, R (2020). An application of simultaneous stochastic optimisation of an open-pit mining complex with tailings management. *International Journal of Mining, Reclamation and Environment* 34(8): 592-607.

Salmi, E, Mousavi, A, Usami, K, and Sellers, E (2022). Value chain optimization for grade engineering by mine-wide frag-mentation optimisation. In: *Wang, X G (eds). Rock fragmentation by blasting* Beijing, China.

Salmi, E F and Sellers, E J (2021). A review of the methods to incorporate the geological and geotechnical characteristics of rock masses in blastability assessments for selective blast design. *Engineering Geology* 281(2): 105970.

Sari, Y A and Kumral, M (2018). A landfill based approach to surface mine design. *Journal of Central South University* 25(1): 159-168.

Sawatsky, L, Mckenna, G, Keys, M-J, and Long, D (2000). Towards minimising the long-term liability of reclaimed mine sites. In: *Haigh, M (eds). Reclaimed land : Erosion control, soils and ecology*. Brookfield: A.A. Balkema, Rotterdam.

Schwartz, E S (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance* 52(3): 923-973.

Scott, M C (2013). Evaluation of energy-efficiency, emission pricing and preconcentration for the optimised development of a au-cu deposit. *Ph.D.*, University of Queensland.

Senécal, R and Dimitrakopoulos, R (2020). Long-term mine production scheduling with multiple processing destinations under mineral supply uncertainty, based on multi-neighbourhood tabu search. *International Journal of Mining, Reclamation and Environment* 34(7): 459-475.

Shishvan, M S and Benndorf, J (2016). The effect of geological uncertainty on achieving short-term targets: A quantitative approach using stochastic process simulation. *Journal of the Southern African Institute of Mining and Metallurgy* 116: 259-264.

Shishvan, M S and Sattarvand, J (2015). Long term production planning of open pit mines by ant colony optimization. *European Journal of Operational Research* 240(3): 825-836.

Silver, D, Huang, A, Maddison, C J, 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, and Hassabis, D (2016). Mastering the game of go with deep neural networks and tree search. *Nature* 529(7587): 484-489.

Silver, D, Schrittwieser, J, Simonyan, K, Antonoglou, I, Huang, A, Guez, A, Hubert, T, Baker, L, Lai, M, and Bolton, A (2017). Mastering the game of go without human knowledge. *Nature* 550(7676): 354-359.

Skousen, J, Rose, A, Geidel, G, Foreman, J, Evans, R, and Hellier, W (1998). Handbook of technologies for avoidance and remediation of acid mine drainage. *National Mine Land Reclamation Center, Morgantown* 131.

Skousen, J, Zipper, C E, Rose, A, Ziemkiewicz, P F, Nairn, R, Mcdonald, L M, and Kleinmann, R L (2017). Review of passive systems for acid mine drainage treatment. *Mine Water and the Environment* 36(1): 133-153.

Smith, M L (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(4): 149-155.

Soltani, S and Safa, M (2015). Optimally locating additional drill holes to increase the accuracy of ore/waste classification. *Mining Technology* 124(4): 213-221.

Stone, P, Froyland, G, Menabde, M, Law, B, Pasyar, R, and Monkhouse, P (2005). Blasor—blended iron ore mine planning optimisation at yandi, western australia. In: *Dimitrakopoulos, R (eds). Orebody modelling and strategic mine planning*. The Australasian Institute of Mining and Metallurgy, Carlton, Vic,. 14: 133-136.

Stone, P, Froyland, G, Menabde, M, Law, B, Pasyar, R, and Monkhouse, P (2018). Blasor—blended iron ore mine planning optimisation at yandi, western australia. *Advances in Applied Strategic Mine Planning*: 39-46.

Storn, R and Price, K (1997). Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization* 11(4): 341.

Straker, J, Mcconnachie, J, Baker, T, Mckenna, G, and Antill, T (2020). Guiding principles for successful mine reclamation in british columbia. In: *British Columbia Mine Reclamation Symposium*, Online.

Strebelle, S (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology* 34(1): 1-21.

Strebelle, S and Cavelius, C (2014). Solving speed and memory issues in multiple-point statistics simulation program snesim. *Mathematical Geosciences* 46(2): 171-186.

Sutton, R S and Barto, A G (2018). Reinforcement learning: An introduction. MIT press, Cambridge, Massachusetts.

Sutton, R S, Mcallester, D, Singh, S, and Mansour, Y (1999). Policy gradient methods for reinforcement learning with function approximation. In: *Solla, S A, Leen, T K and Mulller, K-R (eds). Advances in neural information processing systems 12*. MIT Press, Cambridge, Massachusetts.

Thompson, F, De Oliveira, B C, Cordeiro, M C, Masi, B P, Rangel, T P, Paz, P, Freitas, T, Lopes, G, Silva, B S, and Cabral, A S (2020). Severe impacts of the brumadinho dam failure (minas gerais, brazil) on the water quality of the paraopeba river. *Science of the total environment* 705: 135914.

Upadhyay, S P and 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(2): 153-166.

Upadhyay, S P and Askari-Nasab, H (2019). Dynamic shovel allocation approach to short-term production planning in open-pit mines. *International Journal of Mining, Reclamation and Environment* 33(1): 1-20.

Urbaez, E and Dagdelen, K (1999). Implementation of linear programming model for optimum open pit production scheduling problems. In: *Transactions Society for Mining, Metallurgy, and Exploration*.

Vallee, M (2000). Mineral resource + engineering, economic and legal feasibility = ore reserve. *CIM bulletin* 93(1038): 53-61.

Vargas-Guzmán, J A and Dimitrakopoulos, R (2003). Successive nonparametric estimation of conditional distributions. *Mathematical Geology* 35(1): 39-52.

Vaziri, V, Sayadi, A R, Mousavi, A, Parbhakar-Fox, A, and Monjezi, M (2021). Mathematical modeling for optimized mine waste rock disposal: Establishing more effective acid rock drainage management. *Journal of Cleaner Production* 288: 125124.

Vaziri, V, Sayadi, A R, Parbhakar-Fox, A, Mousavi, A, and Monjezi, M (2022). Improved mine waste dump planning through integration of geochemical and mineralogical data and mixed integer programming: Reducing acid rock generation from mine waste. *Journal of Environmental Management* 309: 114712.

Wambeke, T and Benndorf, J (2017). A simulation-based geostatistical approach to real-time reconciliation of the grade control model. *Mathematical Geosciences* 49(1): 1-37.

Wambeke, T, Elder, D, Miller, A, Benndorf, J, and 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(3): 115-130.

Whittle, G (2007). Global asset optimization. In: *Dimitrakopoulos, R (eds). Orebody modelling and strategic mine planning: Uncertainty and risk management models*. AusIMM Spectrum Series: 331-336.

Whittle, G and Burks, S (2010). Simultaneous mining and mineral processing enterprise optimization for the platinum industry. In: *Proceedings of the 4th International Platinum Conference, Platinum in transition "Boom or Bust", Johannesburg, South Africa*.

Whittle, J (2018). The global optimiser works—what next? In: *Advances in Applied Strategic Mine Planning*, Springer International Publishing, Cham.

Whittle, J and Whittle, G (2007). Global long-term optimisation of very large mining complexes. In: *Application of computers and operations research in the mineral industry: APCOM 2007*, Santiago, Chile.

Wilke, F L and Reimer, T (1977). Optimizing the short term production schedule for an open pit iron ore mining operation. In: *15th APCOM Symposium*, Brisbane, Australia.

Wu, J, Zhang, T, and Journel, A (2008). Fast filtersim simulation with score-based distance. *Mathematical Geosciences* 40(7): 773-788.

Yao, L, Dimitrakopoulos, R, and Gamache, M (2018). A new computational model of high-order stochastic simulation based on spatial legendre moments. *Math Geosci* 50(8): 929-960.

Yokota, F and Thompson, K M (2004). Value of information literature analysis: A review of applications in health risk management. *Medical Decision Making* 24(3): 287-298.

Zhang, J and Dimitrakopoulos, R G (2017). A dynamic-material-value-based decomposition method for optimizing a mineral value chain with uncertainty. *European Journal of Operational Research* 258(2): 617-625.

Zhang, J and Dimitrakopoulos, R G (2018). Stochastic optimization for a mineral value chain with nonlinear recovery and forward contracts. *Journal of the Operational Research Society* 69(6): 864-875.

Zhang, T, Switzer, P, and Journel, A (2006). Filter-based classification of training image patterns for spatial simulation. *Mathematical Geology* 38(1): 63-80.

Zipper, C E, Burger, J A, Skousen, J G, Angel, P N, Barton, C D, Davis, V, and Franklin, J A (2011). Restoring forests and associated ecosystem services on appalachian coal surface mines. *Environmental Management* 47(5): 751-765.