Updating short-term destination policies with new information in a mining complex: A fast mechanism using machine learning

> A. Kumar, R. Dimitrakopoulos, M. Maulen G–2018–82

October 2018

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

**Citation suggérée:** A. Kumar, R. Dimitrakopoulos, M. Maulen (Octobre 2018). Updating short-term destination policies with new information in a mining complex: A fast mechanism using machine learning, Rapport technique, Les Cahiers du GERAD G–2018–82, GERAD, HEC Montréal, Canada.

**Avant de citer ce rapport technique**, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2018-82) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2018 – Bibliothèque et Archives Canada, 2018

> GERAD HEC Montréal 3000, chemin de la Côte-Sainte-Catherine Montréal (Québec) Canada H3T 2A7

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

Suggested citation: A. Kumar, R. Dimitrakopoulos, M. Maulen (October 2018). Updating short-term destination policies with new information in a mining complex: A fast mechanism using machine learning, Technical report, Les Cahiers du GERAD G-2018-82, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (https:// www.gerad.ca/en/papers/G-2018-82) to update your reference data, if it has been published in a scientific journal.

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2018 – Library and Archives Canada, 2018

Tél.: 514 340-6053 Téléc.: 514 340-5665 info@gerad.ca www.gerad.ca

# Updating short-term destination policies with new information in a mining complex: A fast mechanism using machine learning

# Ashish Kumar<sup>*a,c*</sup> Roussos Dimitrakopoulos<sup>*a,b,c*</sup> Marco Maulen<sup>*d*</sup>

<sup>a</sup> GERAD, Montréal (Québec), Canada, H3T 2A7

<sup>b</sup> COSMO–Stochastic Mine Planning Laboratory, McGill University, Montréal (Québec) Canada, H3A 0E8

<sup>c</sup> Department of Mining and Materials Engineering, McGill University, Montréal (Québec), Canada, H3A 0E8

<sup>d</sup> BHP, Technical Manager, Mine Planning, Chile

ashish.kumar@mcgill.ca roussos.dimitrakopoulos@mcgill.ca marco.maulen1@bhpbilliton.com

October 2018 Les Cahiers du GERAD G-2018-82 Copyright © 2018 GERAD, Kumar, Dimitrakopoulos, Maulen

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;

• Peuvent distribuer gratuitement l'URL identifiant la publication. Si vous pensez que ce document enfreint le droit d'auteur, contacteznous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande. The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profitmaking activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Abstract:** New digital technologies including the development of advanced sensors and monitoring devices have enabled a mining complex to acquire new information about the performance of its different components. The short-term production plan, which determines production decisions about extraction sequence, destination policies, fleet assignment and allocation, and processing stream utilization is significantly dependent on the performance of and interaction between the different components of the mining complex. A minor change in the performance of any component can significantly change the processing capabilities and consequently the net revenue of a mining complex. Existing technologies, though they integrate conventionally collected new information, are not able to integrate the new incoming information to adapt the short-term production plan. This paper presents a new continuous updating framework based on policy gradient reinforcement learning and ensemble Kalman filter to adapt the short-term production plan regarding destination policies of material in a mining complex with new information. The framework first uses ensemble Kalman filter to update the uncertainty models of the different components of a mining complex with new information. Then, the updated uncertainty models are utilized in a neural network based policy gradient reinforcement learning algorithm to adapt the short-term destination policies of material in a mining complex. The proposed framework is applied to a copper-gold mining complex, which shows its ability to adapt the short-term destination policies of material with new information. The framework better meets the different production targets while improving the cumulative cash flow compared to the industry standard fixed cut-off grade policy.

**Keywords:** Mining complexes, ensemble Kalman Filter, reinforcement learning, updating destination policies, new information, updating stochastic simulations

# 1 Introduction

A mining complex is an integrated business where the material extracted from a group of mines is sent to different processing streams to produce sellable products. The components of a mining complex include mineral deposits, loading and hauling equipment, and processing streams such as stockpiles, crushers, waste dumps, leach pads, processing mills, and means of transportation, that are interlinked to form a mineral value chain. Uncertainty is a characteristic of a mining complex, starting from the supply of different types of materials extracted from the mineral deposits involved. A long-term production plan of a mining complex determines the annual strategic decisions that maximize net present value (NPV) and meets different production targets, while accounting for uncertainty in the supply of different types of materials (Goodfellow and Dimitrakopoulos, 2016, 2017; Montiel and Dimitrakopoulos, 2015, 2017). The short-term production plan determines the daily/weekly/monthly production decisions within the long-term production plan to meet annual targets. The short-term production plan accounts for uncertainty in the performance of equipment, in addition to supply uncertainty, to determine the production decisions about the extraction sequence, equipment assignment, and allocation (Matamoros and Dimitrakopoulos, 2016; Quigley, 2016). A major short-term production decision is to determine the destination of extracted material in a mining complex and is referred to as destination policy (Asad et al., 2016). The destination decision of extracted materials is significantly dependent on the performance of and interaction between the different components of a mining complex and, therefore, must be adapted with the new incoming information about the performance of the related components.

New digital technologies, including the development of advanced sensors and monitoring devices, have enabled the acquisition of new information about the performance of the different components of a mining complex. Sensors installed on shovels, trucks, conveyor belts, crushers, and processing mills (Dalm et al., 2014, 2017; Goetz et al., 2009; Iyakwari et al., 2016) measure the performance of the mining equipment and processing streams, as well as different pertinent properties of the material being handled. In addition to the new sensor information, conventional sources of new information include blasthole sampling that determines the performance of equipment (Koellner et al., 2004), and tracking devices that track the location of materials (Rosa et al., 2007).

Existing technologies, though they integrate the conventionally collected new information such as blasthole data for grade control and the monitoring of equipment for assignment and dispatch decisions (Kargupta et al., 2010; Nguyen and Bui, 2015), they are not able to integrate the new information and adapt the shortterm production plan accordingly. A continuous updating framework as shown in Figure 1 is needed to adapt and update the short-term production plan of a mining complex with new information generated from both sensors and conventional sources. The continuous updating framework consists of two parts. First, the new information generated from the different sources is used to update the performance of different components of a mining complex, including uncertainty in the supply of materials from the mineral deposits, performance of equipment, and processing streams. Second, the updated uncertainty models of the different components of a mining complex are used to update and adapt the short-term production plan. The updated short-term production plan is fed back to the mining complex to generate updated production forecasts. Benndorf and Buxton (2016) propose a similar framework to update the mine planning decisions with new information. Related is also the work of Hou et al. (2015) and Shirangi (2017) who propose a continuous updating framework to update the production plan of smart oil fields. However, the existing frameworks, both in mine planning and smart oil field, require re-optimization of the production plan, which is computationally expensive with the available optimization techniques. Lamphari (2017) provides a detailed review of the different techniques used for production planning in mining complexes and smart oil fields.

The new information generated in a mining complex can be categorized as "soft" and "hard", based on the precision of their measurement. Sensor-generated information is soft information in the sense as when compared to blasthole and exploration data, which is subtantially more precise and considered as hard data. Therefore, the first part of the continuous updating framework in a mining complex, as shown in Figure 1, aims at generating updated uncertainty models of the different components of a mining complex



Figure 1: A continuous updating framework to update and adapt the short-term production plan of a mining complex with new incoming information

that are consistent with the hard data and minimize the mismatch between the (a) observed and forecasted production data, as well as (b) soft and hard data. Evensen et al. (1994) propose the ensemble Kalman filter (EnKF) that updates the non-linear processes with new information and has long been used for petroleum reservoir flow simulation and production forecasting (Dovera and Della Rossa, 2011; Xue and Zhang, 2014). The ensemble Kalman filter is a two-step assimilation process that first generates a model-based prediction based on initial simulations for a non-linear process and then corrects such predictions with new observed information. The method has been successfully applied to update pertinent attributes of mineral deposits (Benndorf, 2015; Yüksel et al., 2017). Methods such as randomized maximum likelihood (Chen and Oliver, 2012; Sarma et al., 2006; Shirangi, 2017; Vo and Durlofsky, 2014) and Markov mesh models (Panzeri et al., 2016) are also used to update the pertinent petroleum reservoir related attributes. Vargas-Guzmán and Dimitrakopoulos (2002) and Jewbali and Dimitrakopoulos (2011) propose a column-wise decomposition of the covariance matrix (CSSR) to update the pertinent attributes of mineral deposits with new hard data. However, the method cannot integrate the soft information generated from sensors. The outlined methods for updating pertinent attributes of mineral deposits with EnKF and CSSR are limited to a single attribute.

The second part of the updating framework (Figure 1) aims at adapting the short-term production plan of a mining complex with the updated uncertainty models of its different components. Reinforcement learning methods are efficient and fast at adapting decisions to new information. In recent years, reinforcement learning based methods have shown exceptional performance at generating neural network agents that are capable of making very efficient decisions for different complex environments (Mnih et al., 2013; Silver et al., 2016). Paduraru and Dimitrakopoulos (2017) propose Bayesian reinforcement learning to optimize the destination policies of materials in a mining complex. However, the method developed requires a predefined extraction sequence to calculate the expected a posteriori improvement in the objective function during the optimization. Paduraru and Dimitrakopoulos (2018) propose policy gradient reinforcement learning to optimize the neural network destination policies of materials in a mining complex while accounting for supply and equipment performance uncertainty. The neural network destination policies increased the expected NPV by 6.5 % compared to an optimized cut-off grade policy for a copper mining complex. However, the method is (a) limited to a single product mining complex, and (b) does not provide a continuous updating of the short-term production plan regarding destination policies of materials with the new information generated from sensors and conventional sources.

The work presented herein contributes the following. First, it extends the model presented in Benndorf (2015) to updated multiple pertinent attributes in a mineral deposit with new incoming information. Second,

it generalizes the model presented in Paduraru and Dimitrakopoulos (2018) to account for multiple products in a mining complex. Finally, it proposes a continuous updating framework that uses the first and second part to adapt the short-term production plan regarding destiantion policies of mateiral in a multiple product mining complex with new incoming information. In the following sections, first, the proposed continuous updating framework that adapts the short-term production plan in terms of destination policies of material with new incoming information is detailed. Next, an application of the proposed continuous updating framework at a copper-gold mining complex is presented to show the efficiency of the proposed framework compared to the optimized cut-off grade policy. Conclusions and directions for future research follow.

# 2 Method

This section outlines the framework proposed to update the short-term destination policies of material in a mining complex with new incoming information. First, the complete updating framework is provided, then the algorithm related to different parts of the updating framework is detailed.

#### 2.1 Updating framework

Figure 2 presents the proposed framework to update the short-term destination decisions of material in a mining complex with new incoming information while accounting for uncertainty in the supply of material and performance of equipment. The proposed framework to update the short-term destination policies of material in a mining complex consists of two main parts as shown in Figure 2. In the first part (see Sect. 2.2), the initial stochastic simulations of pertinent attributes in the mineral deposits and pertinent attributes related to performance of equipment are updated with the new incoming information. Updating the simulations of attributes related to performance of equipment is straightforward and will not be discussed in this paper. The method to update stochastic simulations of mineral deposits attributes is based on ensemble Kalman filter. In the second part (see Section 2.3), the updated simulations of pertinent attributes of mineral deposit and equipment performance are fed to a stochastic model of a mining complex, which simulates the extraction and hauling of material. The information from the previous step is fed to a neural network and trained using policy gradient reinforcement learning to update the short-term destination decisions of the material. The material from these destinations flows through a set of processing streams, which finally produces the different products in the mining complex. In what follows, a detailed description of the two parts of the updating framework is provided.



Figure 2: Proposed framework for updating short-term destination policies with new information in a mining complex

#### 2.2 Updating stochastic orebody simulations

The method proposed to update simulations of a mineral deposit with new information uses ensemble Kalman filter, which is modified to account for multiple correlated elements in the mineral deposits. The simulations of mineral deposits are herein referred to as ensembles. The complete process to update ensembles with multiple correlated elements based on new information is shown in Figure 3. First, the exploration drill information with multiple elements is de-correlated using minimum/maximum autocorrelation factor (MAF). The decorrelated MAF factors are then used to generate initial ensembles. The new information acquired in the mining complex about the quality of the material is de-correlated using MAF. Then, the decorrelated new information and the initial ensembles are used in the ensemble Kalman filter (EnKF) method to generate the updated ensembles of multiple correlated elements. The updated ensembles are finally transformed back from MAF factors into correlated elements and averaged to mining block sizes that represent the selectivity of the operation in the mining complex.



Figure 3: Updating stochastic simulations of mineral deposits with new information

#### 2.2.1 Notation

- SSet of stochastic orebody simulations,  $\mathfrak{s} \in S$
- Set of mining blocks in a mine m, b in BB
- $\mathcal{N}$ Set of internal nodes in a mineral deposit
- Location of internal node i in a mineral deposit  $x_i$
- VSet of internal nodes  $x_i$  in a mining block b
- ESet of elements in a mineral deposit, e in E
- $F_{e}^{t,\mathfrak{s}}(x_{i})$ MAF value at location  $x_i$  for element e in period t and scenario  $\mathfrak{s}$
- $d_{e}^{t,\mathfrak{s}}\left(x_{i}\right)$ Data value at location  $x_i$  for element e in period t and scenario  $\mathfrak{s}$ 
  - MAF transformation function for element  $e, \mathbb{Z}_e^{t,\mathfrak{s}}(x_i) = \Phi_M^e\left(d_e^{t,\mathfrak{s}}(x_i)\right)$  $\Phi^e_M$
  - $\Phi^e_{\underset{1}{G}}$ Gaussian transformation function for element e
  - $\Phi^{e^{-1}}_{\underline{M}_{-1}}$ Maf inverse transformation function for element e
  - $\Phi^e_G$ Gaussian inverse transformation function for element e
  - Matrix defining contribution of internal nodes towards new information A  $A^T$ Transpose of matrix A
- $\mathbb{Z}_{e}^{t,\mathfrak{s}}\left(x
  ight)$ Vector of  $\mathbb{Z}_e^{t,\mathfrak{s}}(x_i)$  for element e in period t and scneario  $\mathfrak{s}$ 
  - $P_e^{l_e^t}$ MAF value of new information for element e in period t
  - Model based prediction for element e in period t and scenario  $\mathfrak{s}$
  - $\epsilon_e^t$ Noise in the new information for element e in period t
  - Model error covariance matrix for element e in period t
  - Measurement error covariance matrix for element e in period t
- $\begin{array}{c} C^t_{u_e u_e} \\ C^t_{o_e o_e} \\ K^t_e \end{array}$ Kalman gain for element e in period t
- $d_{e}^{t,\mathfrak{s}}\left(b\right)$ Data value of mining block b for element e in period t and scenario  $\mathfrak{s}$

#### 2.2.2 Updating algorithm

A mineral deposit is discretized into an array of three-dimensional volume referred to as mining blocks. The mining blocks are further discretized into multiple internal nodes. Let  $\mathbb{Z}_{e}^{t,\mathfrak{s}}(x)$  be a realization  $\mathfrak{s} \in S$  of the vector of the spatial random field consisting of elements  $\mathcal{F}_{e}^{t,\mathfrak{s}}(x_{i})$ .  $\mathcal{F}_{e}^{t,\mathfrak{s}}(x_{i})$  represents the simulated MAF value of element e at location  $x_{i}$ , at time t, under scenarios  $\mathfrak{s}$ , with  $i \in [1, \mathcal{N}]$ , being the index of internal nodes. Initial ensembles of MAF values are represented by  $\mathbb{Z}_{e}^{t,\mathfrak{s}}(x)$  for the multiple elements in the mineral deposit. Further, let matrix A describes the contribution of each internal nodes at location  $x_{i}$  towards the new information observed in the mining complex. The new information observed in period t is also decorrelated using MAF into MAF factor  $l_{e}^{t}$  for element e. The Gaussian assumptions in ensemble Kalman filter is handled by transforming  $\mathbb{Z}_{e}^{t,\mathfrak{s}}(x)$ ,  $\mathcal{F}_{e}^{t,\mathfrak{s}}(x_{i})$  and  $l_{e}^{t}$  using Gaussian anamorphoses function  $\Phi_{G}^{e}$ . The transformed vectors,  $U_{e}^{t,\mathfrak{s}}(x) = \Phi_{G}^{e}(\mathbb{Z}_{e}^{t,\mathfrak{s}}(x))$ ,  $u_{e}^{t,\mathfrak{s}}(x_{i}) = \Phi_{G}^{e}(\mathcal{F}_{e}^{t,s}(x_{i}))$ , and  $m_{e}^{t} = \Phi_{G}^{e}(l_{e}^{t})$  are then used in the EnKF updating process. A random noise  $\epsilon_{e}^{t}$  is added in the new information to represent the noise with the measurement of new information (Equation 1). The model-based prediction,  $\mathcal{P}_{e}^{t,\mathfrak{s}}$  represents the predictions based on initial ensembles at the location of observed information (Equation 2).

$$o_e^t = m_e^t + \epsilon_e^t \qquad \qquad \forall e \in E \tag{1}$$

$$P_e^{t,\mathfrak{s}} = A_t \cdot U_e^{t,\mathfrak{s}}\left(x\right) \qquad \qquad \forall e \in E, \mathfrak{s} \in \mathbf{S}$$

$$(2)$$

$$U_e^{t+1,\mathfrak{s}}(x) = U_e^{t,\mathfrak{s}}(x) + K_e^t \cdot \left(o_e^t - P_e^{t,\mathfrak{s}}\right) \qquad \forall e \in E, \mathfrak{s} \in S$$
(3)

$$K_e^t = \left(A_t^T \cdot C_{u_e u_e}^t \cdot A_t + C_{o_e o_e}^t\right)^{-1} A_t^T \cdot C_{u_e u_e}^t \qquad \forall e \in E$$

$$\tag{4}$$

EnKF sequentially updates the state of a Gaussian stochastic process based on noisy observations. EnKF uses Equation 3 to update the initial ensembles with the new information based on the Kalman gain. The Kalman gain is calculated using Equation 4 and defines the significance of old information compared to the new information. For instance, if the Kalman gain is small, meaning the new information is inaccurate, then the new information is slightly considered to update the initial ensembles. On the other hand, if the Kalman gain is large, meaning the new information is accurate, then the initial ensembles are updated significantly with the new information.

$$C_{u_e u_e}^{t+1}(x_i) \cong \frac{1}{S} \sum_{\mathfrak{s}=1}^{S} \left( u_e^{t,\mathfrak{s}}(x_i) - \overline{u_e^{t,\mathfrak{s}}(x)} \right) \cdot \left( u_e^{t,\mathfrak{s}}(x_i) - \overline{u_e^{t,\mathfrak{s}}(x)} \right)^T \qquad \forall i \in \mathcal{N}, e \in E$$
(5)

EnKF approximates the model error covariance matrix with a finite set of ensembles (Equation 5). The measurement error covariance matrix  $C_{o_e o_e}^t$  is initialized randomly from a standard normal distribution. The updated ensemble values are back transformed using Gaussian inverse transformation function  $\Phi_G^{e^{-1}}(U_e^{t+1,\mathfrak{s}}(x))$  to generate updated MAF ensemble values  $\mathbb{Z}_e^{t+1,\mathfrak{s}}(x)$ . The updated MAF ensemble values are further back transformed using the MAF inverse transformation function and averaged to generate values of different elements in the mining blocks for different ensembles (Equation 6).

$$d_e^{t+1,\mathfrak{s}}(b) \approx \frac{1}{V} \sum_{i=1}^{V} \Phi_M^{e^{-1}} \left( \mathbb{Z}_e^{t+1,\mathfrak{s}}(x_i) \right) \qquad \forall x_i \in b, b \in B, \mathfrak{s} \in S, e \in E$$
(6)

#### 2.3 Updating short-term destination policies in a mining complexe

The method proposed to update the short-term destination polices of material in a multiple product mining complex uses policy gradient reinforcement learning and neural networks, and is based on the work of Paduraru and Dimitrakopoulos (2018). The method accounts for the uncertainty in the supply of different materials and in the performance of the equipment. A short-term stochastic model detailed in Section 2.3.2 is first used to generate the information required to train the neural network adaptive policies. The method to train and adapt neural network policies is presented in Section 2.3.3.

#### 2.3.1 Notation

S	Set of joint uncertainty scenarios, $s \in \mathbb{S}$
T	Time period, $t \in \mathbb{T}$
$\mathcal{M}$	Set of mines in a mining complex
$\mathcal{B}_m$	Set of mining blocks in a mine $m, b \in \mathcal{B}_m$
${\mathcal C}$	Set of crushers in a mining complex
${\mathcal P}$	Set of processing mills in a mining complex
$\mathcal{L}_O$	Set of oxide leach pads in a mining complex
$\mathcal{L}_S$	Set of sulphide leach pads in a mining complex
$\mathcal W$	Set of waste dumps in a mining complex
$\mathbb{P}_T$	Property tonnage that flows in the mining complex
$\mathbb{P}_M$	Set of metals that flows in the mining complex
$g_{a,b,s}$	Grade of metal $a$ in block $b$ in scenario $s$
$m_{b,s}$	Mass of block $b$ in scenario $s$
$v_{a,i,t,s}$	Amount of property $a$ at location $i$ in period $t$ and scenario $s$
$r_{a,i,s}$	Recovery of property $a$ at location $i$ in scenario $s$
$P_{a,i}$	Profit of product $a$ at locaiton $i$
$U_{a,i,t}$	Upper production target for property $a$ at location $i$ in period $t$
$L_{a,i,t}$	Lower production target for property $a$ at location $i$ in period $t$
$C_{a,i}$	Cost of processing material property $a$ at location $i$
$c_{a,i}^+$	Cost of deviation from upper production target $U_{a,i,t}$ at location i
$c_{a,i}^-$	Cost of deviation from lower production target $L_{a,i,t}$ at location i
$d^+_{a,i,t,s}$	Excess from target $U_{a,i,t}$ at location <i>i</i> in period <i>t</i> and scenario <i>s</i>
$d^{-}_{-}$	Shortage from target $L_{a,i,t}$ at location <i>i</i> in period <i>t</i> and scenario <i>s</i>
$T^{h}$ .	Hauling time from mine $m$ to destination $d$ under equipment scenario s
$T^{q}_{m,d,s}$	Oueue time at destination $d$ under scenario s
$T_{l}^{1}d,s$	Leading time with should at mine m under equipment scenario a
$T_{E}^{I_{m,s}}$	Total actuation time from mine m to destinction d in companie a
$I_{m,d,s}^{-}$	Total extraction time from mine $m$ to destination $a$ in scenario $s$
nIter	Number of iteration
n1 	Number of input neurons
nH	Number of hidden neurons
nO	Number of output neurons
$n_j$	Output neuron k
$O_k$	$W_{i} = h_{i} + h_{i$
$w_{ij}$	Weight with arc from input neuron $i$ to model neuron $j$
$SV_i$	Components of input state vector $i$
$w_{jk}$	weight with arc from hidden neuron j to output neuron $\kappa$
$g_i$	Gradient value at iteration $i, i \in [1, nIter]$
$\eta$	Decay rate
0	Smoothing term
$\gamma$	(1) if block is cont to destinction dimensional t
$z_{b,d,t} =$	1, If block $v$ is sent to destination $a$ in period $t$
$\alpha$ $\alpha$ $\subset [0, 1]$	(0, 0) the wave $(0, 0)$ the formula of the form
$y_{a,i,j,t,s} \in [0,1]$	Amount of property a sent from location i to $j \in$ period t, scenario s

#### 2.3.2 A stochastic model of the mining complex

A stochastic model of a mining complex is presented in this section that uses concepts from discrete event simulation, stochastic modelling, and system dynamics to calculate the total extraction time of materials. Consider an illustrative example shown in Figure 4, where the material is first extracted with shovels that have an uncertain performance with regards to productivity, breakdown time, and repair time. Uncertainty scenarios for the shovel performance are generated from historical data. The material extracted is then loaded into trucks and hauled to different destinations. The decision of hauling the material to a destination is based on a destination policy, which, in this work, is based on a neural network. Uncertainty scenarios for truck performance (cycle time) are also generated from historical data. The trucks at different destinations might have waiting time depending on the performance of the destination. For instance, waiting time at crusher will depend on crushing time of the crushers and the amount of material already being crushed. The total extraction time  $T_{m,d,s}^E$  is, therefore, a function of loading time  $T_{m,s}^l$ , hauling time to a destination  $T_{m,d,s}^h$ , and wait time at a destination  $T_{d,s}^q$  calculated using Equation 7.

$$T_{m,d,s}^{E} = f\left(T_{m,s}^{l}, T_{m,d,s}^{h}, T_{d,s}^{q}\right)$$
(7)



Figure 4: An illustrative example of a stochastic model of the mining complex

The material is crushed with a fixed throughput at the crushers. The crushed material is then conveyed to one of the processing mills using a greedy heuristic based on available capacity at the different processing mills. The processing mills recover the metal from the material and generate the multiple products in the mining complex. The recovery of the processing mills is also stochastic and depends on the quality of the feed material. The stochastic scenarios of equipment performance and processing mills recovery is combined with the stochastic simulations of mineral deposits to generate the joint uncertainty scenarios S.

#### 2.3.3 Updating algorithm

The stochastic model of a mining complex simulates the flow of material in the mining complex under the joint uncertainty scenarios S, which is used to train the neural network adaptive policies. Note, the proposed model only decides the destination of material based on multiple elements in a mining complex, given a fixed production plan. The complete training process of the neural network is presented in Figure 5(a). The joint uncertainty scenarios are fed to the stochastic model to perform the extraction and hauling simulations that generates the information about the state  $(SV_i)$  about the quality and quantity of material extracted and hauled from the different mines under joint uncertainty scenarios.  $SV_i$ , represent the input state vector fed to input neurons in the fully connected feed forward neural network. The input to different hidden neurons is calculated using Equation 8. Equation 9 is used to calculate the output of hidden neurons using the rectified linear function. The input to output neurons is then calculated using Equation 10.

input 
$$(h_j) = \sum_{i \in nI} w_{ij}^h SV_i$$
  $\forall j \in nH$  (8)

output  $(h_j) = \max(0, input(h_j))$   $\forall j \in nH$  (9)

input 
$$(o_k) = \sum_{j \in nH} w_{jk}^o * output(h_j)$$
  $\forall k \in nO$  (10)

$$z_{b,d,t} = \frac{e^{\operatorname{input}(o_k)}}{\sum_k e^{\operatorname{input}(o_k)}} \qquad \forall t \in \mathbb{T}, b \in \mathcal{B}_m, \ d \in \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{W}$$
(11)

The output of output neurons is the decisions variable  $z_{b,d,t}$  that determines whether (1) or not (0) a block b is sent to a destination d in a period t and is calculated using Equation 11. Equation 11 also ensures that the blocks are only assigned to one destination. Equations 12 and 13 are then used to calculate the amount of metal and mass respectively at the different destination.

 $b \in \mathcal{B}_m$ 

$$v_{a,i,t,s} = \sum g_{a,b,s} \cdot m_{b,s} \cdot z_{b,d,t} \qquad \forall t \in \mathbb{T}, a \in \mathbb{P}_M, i \in \mathcal{C} \cup \mathcal{L}_S, s \in \mathbb{S}$$
(12)

$$v_{a,i,t,s} = \sum_{b \in \mathcal{B}_m} m_{b,s} \cdot z_{b,d,t} \qquad \forall t \in \mathbb{T}, \ a \in \mathbb{P}_T, i \in \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{W}, \ s \in \mathbb{S}$$
(13)

The material from the different destination  $i \in C \cup \mathcal{L}_S$  is further sent to different processing streams  $j \in \mathcal{P} \cup \mathcal{L}_O$ . Processing stream utilization decisions  $y_{a,i,j,t,s}$  represents the amount of material sent from destination i to j and is decided with a greedy heuristic based on available capacity. Equation 14 is used to calculate the material at the different processing streams in the mining complex. Equation 15 ensures that flow conservation is preserved with the processing stream utilization decisions.

$$v_{a,j,t,s} = \sum_{i \in \mathcal{C}} y_{a,i,j,t,s} \cdot v_{a,i,t,s} \qquad \forall t \in \mathbb{T}, \ a \in \mathbb{P}_M \cup \mathbb{P}_T, j \in \mathcal{P} \cup \mathcal{L}_O, \ s \in \mathbb{S}$$
(14)

$$\sum_{j \in \mathcal{P} \cup \mathcal{L}_{\mathcal{O}}} y_{a,i,j,t,s} = 1 \qquad \forall t \in \mathbb{T}, i \in \mathcal{C}, \ s \in \mathbb{S}$$
(15)

$$\forall t \in \mathbb{T}, a \in \mathbb{P}_M, i \in \mathcal{P} \cup \mathcal{L}_S \cup \mathcal{L}_O, s \in \mathbb{S}$$
(16)

$$v_{a,i,t,s} + d_{a,i,t,s}^{-} \ge L_{a,i,t} \qquad \forall t \in \mathbb{T} , a \in \mathbb{P}_{M}, i \in \mathcal{P} \cup \mathcal{L}_{S} \cup \mathcal{L}_{O}, s \in \mathbb{S} \qquad (17)$$

Equations 16 and 17 are used to calculate the amount of deviation from different production targets in the mining complex. The metal is finally recovered at the different processing destinations. The objective or reward function is given by Equation 18.

$$f(X) = \frac{1}{\|\mathbb{S}\|} \left\{ \sum_{s \in \mathbb{S}} \sum_{t \in \mathbb{T}} \left\{ \sum_{\substack{i \in \mathcal{P} \cup \\ \mathcal{L}_O \cup \mathcal{L}_S}} \sum_{a \in \mathbb{P}_M} P_{a,i} \cdot v_{a,i,t,s} \cdot r_{a,i,s} \right\}$$
Part 1  
$$- \sum_{\substack{i \in \mathcal{P} \cup \mathcal{C} \cup \\ \mathcal{L}_S \cup \mathcal{L}_O \cup \mathbb{M}}} \sum_{a \in \mathbb{P}_T} C_{a,i} \cdot v_{a,i,t,s}$$
Part 2  
$$- \sum_{\substack{i \in \mathcal{P} \cup \\ \mathcal{L}_S \cup \mathcal{L}_O \cup \mathbb{M}}} \sum_{a \in \mathbb{P}_M} (c_{a,i}^+ \cdot d_{a,i,t,s}^+ + c_{a,i}^- \cdot d_{a,i,t,s}^-) \right\}$$
Part 3

Part I in the objective function represents the profits from selling different products, Part II represents the different costs incurred throughout the flow of material, and Part III represents the penalties incurred due to deviation from different production targets. The objective function is an expected value and is not necessarily differentiable to be used in gradient ascent process of neural network training. However, policy gradient reinforcement learning offers the ability that given a reward function f and probability density function  $z_W$  parameterized by W, the equality in Equation 19 below holds true.

$$\nabla_W E_{x \sim z_W(x)} \left[ f(x) \right] = E_{x \sim z_W(x)} \left[ f(x) \nabla_W \log \left( z_W(x) \right) \right]$$
(19)

f(x) in Equation 19 corresponds to the reward function and  $z_W(x)$  corresponds to the action-selection probabilities computed using Equation 11. The weight matrix W contains the values of the hidden  $w_{ij}^h$  and the output neurons  $w_{jk}^o$ . As its common in stochastic gradient method,  $E_{x\sim z_W(x)}[f(x)\nabla_W \log(z_W(x))]$ is replaced with  $f(X)\nabla_W \log(z_W(X))$ , where f(X) represent the cumulative reward obtained during the planning horizon T using the vector of decisions X. The gradient of  $\log(z_W(X))$  can, therefore, be calculated using Equation 20, where the sum is over the planning horizon and over all the destinations. Finally, the stochastic approximation of  $\nabla_W E_{x \sim z_W(x)} [f(x)]$  can be computed using Equations 18–20.

$$\nabla_{W} \log \left( z_{W} \left( X \right) \right) = \sum_{t \in \mathbb{T}} \sum_{d \in \mathcal{C} \cup \mathcal{L}_{S} \cup \mathcal{W}} \nabla_{W} \log z_{W}(d) z_{b,d,t}$$

$$\tag{20}$$

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

$$(21)$$

$$W_{i+1} = W_i + \frac{\eta \nabla_W E_{x \sim z_W(x)} \left[ f(x) \right]}{\sqrt{g_{i+1}} + \partial} \qquad \forall i \in [1, nIter]$$
(22)

The weight matrix  $W = \{w_{ij}^h, w_{jk}^o\}$  of the neurons in the neural network is initialized randomly and updated using the gradient ascent method named RMSprop (Hinton et al., 2012) through reinforcement learning. The RMSprop method has better stability and convergence speed and uses Equations 21 and 22 to back propagate and update the weight of the neurons in the training phase of the neural network. The neural network is trained until the pre-defined stopping criteria (*nIter*) is reached.



Figure 5: Process of training the neural network and adapting to new information

The training phase of the neural network allows the generation of destination policies that can adapt to new information. Figure 5(b) represents the process of adapting the neural network destination policies when new information is acquired in a mining complex. The new information is first used to update the joint uncertainty scenarios using the method outlined in Section 2.2, the updated joint uncertainty scenarios are then fed to the stochastic model outlined in Section 2.3.2, which simulates the extraction and hauling of material. The information from the previous step is fed to the trained neural network that decides the destination of material and then a greedy heuristic is used to determine the utilization of the different processing streams. Finally, the forecasts for the different production targets are calculated using Equations 12–18 and further evaluated regarding the probability to meet the different production targets. If the production targets are not met, the neural network is further re-trained using the weight matrix from the previously trained network for very few iterations. The retraining of the neural network for few iterations allows one to adjust the weight of the neural network to meet the different production targets better.

# 3 Application at the escondida mining Complex

The proposed framework for updating short-term destination policies of material is applied at a coppergold mining complex, which demonstrates the useful and applied aspects of the proposed method. In the case study, the new information used to update the stochastic simulations of mineral deposits with multiple elements is the blasthole data collected during the short-term operations. The neural network destination policies accounts for uncertainty in (a) supply of multiple material with multiple elements, (b) performance of equipment related to its availability, cycle times, utilization, downtime, repair time, and productivity, and (c) recovery of metal in processing mills. However, the framework is flexible to include different types of new information related to the supply of material and the performance of different components in the mining complex, in the updating framework.

#### 3.1 Overview of the Escondida mining complex

The Escondida mining complex consists of two mineral deposits (Escondida and Escondida Norte) with 2.32 and 1.84 million mining blocks of size 25x25x15 m<sup>3</sup>. The mineralization has eight different mine zones each for the two deposits. The material is extracted from both the deposits and is sent to one out of seven possible destinations (five crushers, one sulphide leach pad, and one waste dump) as shown in Figure 6. For measuring the performance of the proposed framework, a part of the deposit that consist of 5,581 mining blocks in each deposit extracted over 210 days is used.



Figure 6: Escondida mining complex

Material from five different crushers is then processed at three different processing mills LC, LS, and OGP1, respectively, and an oxide leach pad that supplies material to the port and copper cathode plant. The port and copper cathode plant produce the final product of the mining complex which is transported and sold to different customers and/or the spot market. While the copper cathode plant produces only copper cathodes, the port produces as primary product copper concentrate and secondary products are gold (Au), silver (Ag), and molybdenum (Mo) concentrate. The production targets with different components of the mining complex are presented in Table 1. The operational and economic parameters used in the case study are outlined in Table 2. The economic parameters are scaled for confidentiality purposes.

Table 1: Production targets at Escondida mining complex

Attribute	Target
Crusher 1, Crusher 2, Crusher 3, Crusher 4, and Crusher 5 Mill 1 – LC, Mill 2 – LS, and Mill 3 – OGP1 Oxide leach pad and sulphide leach pad	52055, 10137, 128767, 82192, and 117808 tonne/day 112329, 134247, and 153425 tonne/day 82192 and 369863 tonnes/day
Arsenic grade limit	0.2~%

## 3.2 Base case – fixed cut-off grade policy

The fixed cut-off grade policy is an optimized single or multiple element fixed cut-off value of pertinent attributes in the mineral deposit to decide the destination of the material (Lane, 1984, 1988; Rendu, 2014). The destination policy currently used at the Escondida mining complex is a single element (Copper) fixed cut-off grade policy optimized using Lane's approach (Lane, 1988). The Escondida mining complex is a major

Attribute	Value
Number of schedule mining blocks - Escondida and Escondida Norte	5581 and 5581
Scheduling period	210 days
Recovery of copper at oxide leach pad and sulphide leach pad	0.651  and  0.275
Recovery of copper at processing mills	Stochastic
Recovery of gold, silver, and molybdenum at processing mills	0.25
Processing cost – mills, oxide leach pad, and sulphide leach pad	3.82, 6.4, and 1.84 /tonne
Selling cost – mills, oxide leach pad, and sulphide leach pad	0.26, 0.25, and 0.25 /lb
Selling price – copper, gold, silver, and molybdenum	5511, 35.2x10 <sup>6</sup> , $4.9x10^5$ , and $1.3x10^4$ \$/ tonne
Fixed processing cost – LC, LS, and OGP1	11925.5, 14252.5, and 16288.5  hour
Penalty arsenic grade (As)	3  / $%$ above limit

<b>T</b> 1 1 0	<u> </u>							
Table 7	Unerational	and	economic	narameters t	nr	Fecondida	mining	complex
Tubic 2.	operational	unu	ccononne	purumeters r		Loconalaa		complex

producer of copper products and, therefore, does not consider secondary products such as gold, silver, and molybdenum in the optimization of its fixed cut-off grade policy. The details of the fixed cut-off grade policy are outlined in Table 3. First, the material is classified as sulphide high grade (SHG), sulphide low grade (SLG), oxide and waste depending on the ratio of soluble copper (CuS) to total copper (CuT). The material classification is necessary to determine the possible processing destinations allowed to process the material. For example, SHG material can only be processed at processing mills. However, the decision of whether (1) or not (0) to process the material at an allowed processing destination depends on the fixed cut-off grade mentioned in Table 3. For instance, oxide material can only be processed at an oxide leach pad if the soluble copper is more than 0.2%; otherwise, it is sent to waste dump.

Table 3: Fixed cut-off grade policy used at Escondida mining complex

Material classification	Criteria	Destinations	Fixed cut-off grade policy
SHG	$\frac{CuS}{CuT} \leq 0.2$	Processing Mill	$CuT \geq 0.52\%$
SLG	$\frac{\rm CuS}{\rm CuT} < 0.5$	Sulphide Leach Pad	$CuT \geq 0.33\%$
Oxide	$\frac{\rm CuS}{\rm CuT} \geq 0.5$	Oxide Leach Pad	$CuS \geq 0.2\%$
Waste	Otherwise	Waste Dump	Otherwise

#### 3.3 Adaptive destination policies

The neural network destination policies proposed in this work decide the destination of mining blocks based on the properties of multiple elements in the mining block, as well as the performance of and interaction between the different components in the mining complex. In addition, the proposed method also adapts such destination decisions of mining blocks with incoming new information in the mining complex (See Section 2.3.3). Similar to the fixed cut-off grade policy, the material is first characterized as SHG, SLG, oxide, and waste based on the criteria mentioned in Table 3 to find the allowed processing destinations for a mining block. However, instead of using the fixed cut-off grade policy, the neural network destination policies are used to decide, update, and adapt the destination of material. Three different neural networks are built and trained using policy gradient reinforcement learning mentioned in Section 2.3 that decides whether (1) or not (0) to process the material at (i) processing mills, (ii) a sulphide leach pad, or (iii) an oxide leach pad.

#### 3.4 Parameter selection

This section discusses the selection of different parameters associated with the proposed adaptive neural network destination policies. The state vector information  $SV_i$  consists of 7 to 32 different type of information depending on the complexity of the processing destination, and are fed to the input neurons of the neural network. For instance,  $SV_i$  for the processing mill neural network consists of information about the mass of mining block, different elements such as total copper, soluble copper, arsenic, gold, silver, and molybdenum in the mining block, the material being crushed and leached, the performance of equipment, and the wait

times at the crushers. Similarly, the number of hidden neurons in the neural network ranges from 300 to 800 depending on the number of input neurons. There are only two output neurons to decide whether (1) or not (0) the material is processed at the respective destination. Learning rate and decay the rate with the neural network is set to  $10^{-3}$  and 0.99 respectively as suggested in Hinton et al., (2012). The smoothing term is set to  $10^{-6}$  (Ruder, 2016). The weight of neurons in the neural network is initialized randomly using the Xavier initialization (Glorot and Bengio, 2010). The number of iterations to train the neural network is decided based on the number of simulations of mineral deposits S considered in the joint uncertainty scenarios  $\mathbb{S}$  to train the neural network. The relation between the number of simulations of mineral deposits and the number of iterations is given by Equation 23, and is decided based on the progression of the objective function in the training phase of the neural network with 15 mineral deposits simulations (S). At around 7,500 training iterations, the objective function plateaus for 15 mineral deposits simulations, which justifies the relation in Equation 23.

$$nIter = 500 * S \tag{23}$$



Figure 7: Progression of objective function over iteration

The number of simulations of mineral deposits to use for training the neural network is decided by training the network with different values of S and then testing the trained network on a new set of 100 joint uncertainty scenarios to compare the performance in terms of meeting different production targets. The value of the S at which the solution stabilizes is selected to use for training the neural network destination policies. The details of this study are shown in Table 4. From Table 4, it can be seen that after 25 simulations the number of decisions does not change significantly and also the change in the objective function value is negligible, which highlights that 25 simulations are enough to achieve a stable solution for the proposed short-term adaptive neural network destination policies.

Table 4: Comparison of training the neural network with varying number of mineral deposits simulations

			Sim	ulations	(5)		
Destination	_		SIII	luiations	(3)		
	5	10	15	20	25	30	35
Mill	2040	2008	1967	2053	2155	2070	2095
Sulphide leach pad	3111	3357	3447	3758	3303	3218	3544
Oxide leach pad	308	301	301	294	304	301	308
Waste	4010	3758	3873	3431	3515	3978	3705
Objective Function (%)	110.85	110.67	111.96	113.65	115.32	115.28	115.55

### 3.5 Results

The results of the proposed adaptive neural network destination policies to update short-term destination policies with new information are presented in this section. Results are reported using the 10<sup>th</sup>, 50<sup>th</sup>, and 90<sup>th</sup> percentile risk profiles (P10, P50, and P90 respectively) of the different performance indicators considering 100 joint uncertainty scenarios. The 100 joint uncertainty scenarios used for testing the neural network policies are different from the ones used to train the network. The forecasts of the production targets with

the proposed framework are compared to the forecasts of the fixed cut-off grade policy over the same 100 joint uncertainty scenarios throughout its presentation and discussion, to highlight the differences and added value of the adaptive framework, where appropriate. The training phase of the neural network takes about 52 hours with 12,500 iterations on an Intel processor core if 2,600s with 8GB of RAM. However, it only takes about one hour to update the stochastic simulations of the two mineral deposits and to adapt the destination decisions of the mining blocks for a period of 210 days using the proposed adaptive framework. The results are presented for both the destination policies over initial and update stochastic simulations of mineral deposits to highlight the contribution of the different part of the updating framework.

#### 3.5.1 Updated stochastic simulations of mineral deposit

Figure 8 shows one of the initial and updated simulations of the total copper mineral attribute of the Escondida deposit at block support. Six different correlated elements: soluble copper, total copper, arsenic, gold, silver, and molybdenum, are updated with the new blasthole data. The initial stochastic simulations of six correlated elements in the two mineral deposits, conditional to the exploration drillhole samples, are generated using a generalized sequential Gaussian simulation (Dimitrakopoulos and Luo, 2004). The initial stochastic simulations are updated using the method discussed in Section 2.2 with the new blasthole data collected during the short-term operations. The mineral deposits consist of eight different mine zones each. Therefore, to respect such geological features of the deposits, the blasthole data in a mine zone is only considered to update the mining blocks in the same mine zone. It is clear from Figure 8 that the updated simulations maintain the significant structures inferred from the exploration drillhole data and updates the local characteristics with the new blasthole data. A histogram of the initial and updated simulations at point support confirms such results, where the distribution of total copper in bench 1 for Escondida is very different for the initial and updated simulations. The updated simulations show a higher proportion of high-grade material, as compared to the initial simulations. Figure A1 in Appendix A1 shows the results of updating simulations for bench 2 for Escondida deposit. The number of blast hole data used to update the two deposits is presented in Table A1 in the Appendix A1.



Figure 8: Updated block simulations compared to initial block simulations for bench 1 for Escondida deposit

#### 3.5.2 Production targets

The forecasts of the different production targets mentioned in Table 1 are shown in this section for the adaptive neural network destination policies and compared to the fixed cut-off grade policy. Figure 9(a) shows the risk profile of meeting the capacity target with mill-2 for initial simulations using neural network destination policies compared to the cut-off grade policy in Figure 9(b). The neural network destination

policies are better at meeting the target with maximum utilization of the mill's capacity, as compared to high fluctuations and lower chances of meeting the target in the cut-off grade policy. When the simulations are updated with the new blasthole data, and the two destination policies are used to update the destination decisions, it is observed that the neural network destination policies (Figure 9 (c)) are performing even better than before with an increased chance of meeting production targets compared to the high fluctuations in the cut-off grade policy (Figure 9 (d)). Figure A3 and Figure A4 in the Appendix A2 show the risk profiles of mill-1 and mill-3 capacity targets respectively.



Figure 9: Forecasts of the capacity target of mill-2 with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations

Figure 10(a) and Figure 10(b) show the risk of meeting the blending target of arsenic at mill-2 for initial simulations with adaptive neural network destination policies and fixed cut-off grade policy respectively. The neural network policies have higher chances of meeting such a target with minimal deviations only after 80 days, as compared to the cut-off grade policy, which has higher chances of deviating from such targets, more specifically during first 80 days. After the simulations are updated with blasthole data and the two destination policies are used to adapt the decisions, we see that both the policies are not able to meet the blending restrictions as shown in Figure 10(c) and Figure 10(d). The lower chances of meeting the arsenic target with the updated destination decisions are because the extraction decision (when to extract the mining blocks) is fixed in the proposed framework; thus, the policies are only deciding the destination decisions, it is hard to control the arsenic concertation in the mill without adapting the extraction sequence. Figure A5 and Figure A6 in Appendix A2 show the risk profile of the arsenic blending target with mill-1 and mill-3 respectively.

#### 3.5.3 Metal Production

The results presented for metal production in this section are scaled for confidentiality purposes (cut-off grade policy for initial simulations being 100%). Figure 11(a) and Figure 11(b) represent the risk profile of cumulative copper production at the mills for initial simulations with adaptive neural network destination policies and fixed cut-off grade policy respectively. The neural network destination policies recover 12% higher copper metal (112%), as compared to the cut-off grade policy (100%). The simulations are updated with blasthole data, and the two destination policies are used to update the destination decisions. The neural network destination policies further increase the metal production (119%) by an additional 7 % compared to



Figure 10: Forecasts of arsenic blending target of mill-2 with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations



Figure 11: Forecasts of cumulative copper production at the processing mills with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations

only an 8% increase in the cut-off grade policy, which is still less than the initial metal production from the neural network destination policies (112%).

Figure 12, Figure 13, and Figure 14 show the risk profiles of the production of secondary products gold, silver, and molybdenum concentrate using the neural network destination policies and the fixed cut-off grade policy.

The neural network destination policies generate 23% higher gold (123%) (Figure 12(a)), 29% higher silver (129%) (Figure 13(a)), and 28% higher molybdenum (128%) (Figure 14(a)) products for initial simulations compared to the cut-off grade policy (100%) (Figure 12(b), Figure 13(b), and Figure 14(b)). The adapted

decisions of neural network destination policies further increase the production of secondary products (Figure 12(c), Figure 13(c), and Figure 14(c)) by an additional 27% for gold (150%), 42% for silver (171%), and 48% for molybdenum (176%) compared to an increase of 35% for gold (135%), 56% for silver (156%), and 60% for molybdenum (160%) for cut-off grade policy (Figure 12(d), Figure 13(d), and Figure 14(d)). The updated decisions with the predefined cut-off grade policy still produce significantly less quantity of secondary products compared to the updated decisions from adaptive neural network destination policies.



Figure 12: Forecasts of total gold production at the processing mills with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations



Figure 13: Forecasts of total silver production at the processing mills with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations



Figure 14: Forecasts of total molybdenum production at the processing mills with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations

#### 3.5.4 Cash Flows

The results presented for cumulative cash flows in this section are scaled for confidentiality purposes (cut-off grade policy for initial simulations being 100%). Figure 15 shows the risk profile of cumulative cash flow over 210 days with the adaptive neural network destination policies and fixed cut-off grade policy. The neural network destination policies present a 15% higher cumulative cash flow compared to cut-off grade policy for the initial simulations (Figure 15(a)). Once the simulations are updated with the new blasthole data, and the decisions are adapted using the two destination policies, the neural network destination policies further improve the cumulative cash flow by an additional 7% to result in a value of 122% compared to 111% in the updated decisions with the fixed cut-off grade policy (Figure 15(b)).



Figure 15: Forecasts of the cumulative cash flow of the mining complex with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations

#### 3.5.5 Updated destination decisions

Figure 16 shows the destination of mining blocks for bench 1 in the Escondida deposit for the adaptive neural network destination policies and the fixed cut-off grade policy. Figure 16(a) shows the destination decisions of the neural network destination policies compared to the cut-off grade policy in Figure 16(b) for initial simulations. The adapted destination decisions of the neural network destination policies and the cut-off grade policy are shown in Figure 16(c) and Figure 16(d), respectively. The neural network destination

decisions are very different from the cut-off grade policy decisions for initial and update simulations, which results in the better chances of meeting production targets, consistently higher cumulative cash flows, and increased metal production. The reason for such a difference in the decision of the two destination policies is because of following reasons:

- 1. The ability of the neural network destination policies to account for and capitalize on the performance of and interaction amongst the different components in the mining complex, thus enabling complex decision-making under different sources of uncertainties.
- 2. Integrating multiple sources of uncertainty, such as the supply of material from multiple mines, equipment performance regarding its availability, utilization, cycle times, breakdown time, repair time, productivity, and recovery of metal in the different processing mills, during the decision-making process of neural-network destination policies.
- 3. Accounting for multiple products, such as copper, gold, silver, and molybdenum, as well as deleterious elements such as arsenic while deciding the destination of mining blocks in a mining complex.

In addition, neural network destination policies are well trained and suited to update the destination decisions of mining blocks and to respond to different types of new information acquired during the shortterm operation in a mining complex, which results in better destination decisions of mining blocks that meet the different production targets.



Figure 16: Destination decisions of mining blocks for bench 1 in Escondida deposit with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations

# 4 Conclusions

This paper presents a new framework based on reinforcement learning to adapt and update the short-term destination decisions of material in a mining complex based on the new incoming information. The updating framework consists of two parts: first the new information is used to update the stochastic simulations of mineral deposits with multiple elements, second the updated stochastic simulations of mineral deposits with multiple elements along with the stochastic simulations of equipment performance attributes are used in the reinforcement learning framework to adapt and update the destination decision of mining blocks.

The proposed framework is applied at a copper-gold mining complex and shows its excellent performance to respond better and integrate the new incoming information. In the case study, blasthole data are used as new information to first update the stochastic simulations of mineral deposits with multiple elements. The updated stochastic simulations, along with stochastic simulations of equipment performance attributes, are used to update the destination of mining blocks with neural network destination policies. The results are compared to an industry standard fixed cut-off grade policy. The proposed framework better meets the capacity and blending requirements of the different processing mills in a mining complex compared to the cut-off grade policy. The proposed framework also increases the metal production by 12%, 23%, 29%, and 28% for copper, gold, silver, and molybdenum respectively, as compared to the fixed cut-off grade policy for the initial simulation. The increased metal production results in a 15% higher cash flow for neural network destination policies for initial simulations. The stochastic simulations of mineral deposits are updated with the blasthole data collected during the short-term operations, and then the two destination policies are used to update the destination decisions of mining blocks. The neural network destination policies are better at responding to such new information and adapt the destination decisions more intelligently to meet the targets better. In addition, the updated destination decisions of neural network destination policies increase the metal production by an additional 7%, 27%, 42%, and 48% for copper, gold, silver, and molybdenum, respectively. The increased metal production results in an additional increase of 7% in cash flow for the updated decisions from neural network destination policies. The proposed method only adapts the destination decisions of the mining blocks, thus limiting the full potential and use of new information. In the future, a framework that can adapt the short-term extraction sequence might meet the different production targets more closely through continuously adapting decisions with the new incoming information acquire during the short-term operations.

## Appendix

# A1 Updating stochastic simulations of mineral deposits with blasthole data

This section outlines the additional results of updating the stochastic simulations of mineral deposits with new blasthole data. Table A1 represents the number of blasthole data used to update the two benches in the two mineral deposits in the Escondida mining complex. Figure A1 represents one of the initial and updated stochastic simulations of the copper total mineral attribute in the Escondida deposit for bench 2 at block support with the new blasthole data. It is very clear from the figure that the updated simulations preserve the structures inferred from the exploration drillhole samples and update the local features of the deposits with the new blasthole data. The histograms of initial and updated simulations at point support in Figure A1 also show that the updated simulation has a higher proportion of high-grade material compared to initial simulations.

Danah Manahan	Number of blasthole data			
Bench Number	Escondida	Escondida Norte		
Bench 1	3437	222		
Bench 2	3309	258		

Table A1: Number of blasthole data used to update the stochastic simulations of the two mineral deposits

# A2 Updating short-term destination policies

This section outlines the additional results of updating and adapting destination policies with the proposed framework. Figure A2 represents the destination decisions of mining blocks for bench 2 in the Escondida deposit using the neural network destination policies and cut-off grade policy. Figure A2(a) and Figure A2(b) represent the destination decisions of the neural network destination policies and cut-off grade policy and cut-off grade policy respectively for initial simulations. The updated decisions with neural network destination policies and cut-off grade policy are shown in Figure A2(c) and Figure A2(d) respectively.



Figure A1: Updated block simulations compared to initial block simulations for bench 2 for Escondida deposit



Figure A2: Destination decisions of mining blocks for bench 2 in Escondida deposit with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations

Figure A3(a) and Figure A3(b) shows the risk profiles of the capacity target with mill-1 using neural network and cut-off grade policy for initial simulations. The neural network policies are performing better than the cut-off grade policy to meet the capacity target. Figure A3(c) and Figure A3(d) shows the risk of meeting capacity target with mill-1 when the decisions are adapted with neural network and cut-off grade policy respectively. It is observed that both the policies have lower chances of meeting the mill-1 target arising from combined uncertainty in supply of material from both the mineral deposits. Figure A4(a) and Figure A4(b) represents the risk profiles for mill-3 for both the policies for initial simulations. The neural network policies can meet the target better than the cut-off grade policy, which has high fluctuations and lower chances of meeting the mill-1 target. When the decisions are adapted with the neural network (Figure A4(c)) and the cut-off grade policy (Figure A4(d)), the neural network policies can effectively adapt decisions to such new information and, therefore, increase the chances of meeting the target compared to high fluctuations in the cut-off grade policy.



Figure A3: Forecasts of the capacity target of mill-1 with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations



Figure A4: Forecasts of the capacity target of mill-3 with the neural network destination policies compared to the cut-off grade policy

Figure A5 and Figure A6 show the risk profile of arsenic blending target for mill-1 and mill-3 respectively. It is observed that the destination decisions with the neural network and cut-off grade policies show very few deviations from such blending targets, however, with updated simulations, such targets will be violated, and large deviations are expected with both the policies. Since the extraction sequence is fixed for the proposed work, the model is only deciding the destination of material. Therefore, the high concentration of arsenic in the simulations might result in a high concentration of arsenic at the processing mills.



Figure A5: Forecasts of arsenic blending target of mill-1 with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations



Figure A6: Forecasts of arsenic blending target of mill-3 with the (a) neural network destination policies for initial simulations, (b) cut-off grade policy for initial simulations, (c) neural network destination policies for updated simulations, and (d) cut-off grade policy for updated simulations

# References

Asad, M.W.A., Qureshi, M.A., and Jang, H. (2016). A review of cut-off grade policy models for open pit mining operations. Resources Policy, 49, 142–152

Benndorf, J. (2015). Making use of online production data: Sequential updating of mineral resource models. Mathematical Geosciences, 47(5), 547–563

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

Chen, Y., and Oliver, D.S. (2012). Ensemble randomized maximum likelihood method as an iterative ensemble smoother. Mathematical Geosciences, 44(1), 1-26

Dalm, M., Buxton, M.W.N., Van Ruitenbeek, F.J.A., and Voncken, J.H.L. (2014). Application of near-infrared spectroscopy to sensor based sorting of a porphyry copper ore. Minerals Engineering, 58, 7–16

Dalm, M., Buxton, M.W.N., and van Ruitenbeek, F.J.A. (2017). Discriminating ore and waste in a porphyry copper deposit using short-wavelength infrared (SWIR) hyperspectral imagery. Minerals Engineering, 105, 10–18

Dimitrakopoulos, R., and Luo, X. (2004). Generalized sequential Gaussian simulation on group size v and screen-effect approximations for large field simulations. Mathematical Geology, 36(5), 567–590

Dovera, L., and Della Rossa, E. (2011). Multimodal ensemble Kalman filtering using Gaussian mixture models. Computational Geosciences, 15(2), 307–323

Evensen, G., Carlo, M., and Carlo, M. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research: Oceans, 99(C5), 10143–10162

Glorot, X., and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In Proceedings: International Conference on Artificial Intelligence and Statistics, 249–256

Goetz, A.F.H., Curtiss, B., and Shiley, D.A. (2009). Rapid gangue mineral concentration measurement over conveyors by NIR reflectance spectroscopy. Minerals Engineering, 22(5), 490–499

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

Goodfellow, R., and Dimitrakopoulos, R. (2017). Simultaneous Stochastic Optimization of Mining Complexes and Mineral Value Chains. Mathematical Geosciences, 49(3), 341–360

Hinton, G.E., Srivastava, N., and Swersky, K. (2012). Neural Netwrok for machine Learning Lecture 6a - Overview of mini-batch gradient descent. http://www.cs.toronto.edu/~{}tijmen/csc321/slides/lecture\_slides\_lec6.pdf, Toronto

Hou, J., Zhou, K., Zhang, X.S., Kang, X.D., and Xie, H. (2015). A review of closed-loop reservoir management. Petroleum Science, 12(1), 114–128

Iyakwari, S., Glass, H.J., Rollinson, G.K., and Kowalczuk, P.B. (2016). Application of near infrared sensors to preconcentration of hydrothermally-formed copper ore. Minerals Engineering, 85, 148–167

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

Kargupta, H., Srakar, K., and Gilligan, M. (2010). MineFleet (R): An overview of a widely adopted distributed vehicle performance data mining system. In Proceedings: International Conference on Knowledge Discovery and Data Mining, ACM SIGKDD, 37–46

Koellner, W.G., Brown, G.M., Rodríguez, J., Pontt, J., Cortés, P., and Miranda, H. (2004). Recent advances in mining haul trucks. IEEE Transactions on Industrial Electronics, 51(2), 321–329

Lamphari, A. (2017). Mine planning and oil field development: A survey and research potentials. Mathematical Geosciences, 49(3), 395-437

Lane, K.F. (1984). Cutoff grades for two minerals. In Proceedings: International Symposium on Application of Computers and Operations Research in Minerals Industries, 485–492

Lane, K.F. (1988). The economic definition of ore: Cut-off grades in theory and practice. Mining Journal Books, London

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

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. (2013). Playing Atari with deep reinforcement learning. ArXiv, 1312.5602, 1–9

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

Nguyen, D., and Bui, X. (2015). A real-time regulation model in multi-agent decision support system for open pit mining. In Proceedings: International Symposium Continuous Surface Mining-Aachen, Springer, Cham, 255–262

Paduraru, C., and Dimitrakopoulos, R. (2017). Adaptive policies for short-term material flow optimization in a mining complex. Mining Technology, 6668, 1–8

Paduraru, C., and Dimitrakopoulos, R. (2018). Responding to new information in a mining complex : Fast mechanisms using machine learning. (Submitted), 1–30

Panzeri, M., Della Rossa, E.L., Dovera, L., Riva, M., and Guadagnini, A. (2016). Integration of Markov mesh models and data assimilation techniques in complex reservoirs. Computational Geosciences, 20(3), 637–653

Quigley, M. (2016). The stochastic optimization of long and short-term mine production schedules incorporating uncertainty in geology and equipment performance, Masters Thesis. McGill University

Rendu, J.-M. (2014). An introduction to cut-off grade estimation. Society for Mining, Metallurgy & Exploration (2nd ed.), Englewood, Colarado

Rosa, L., David, Valery, W., Wortley, M., Ozkocak, T., and Pike, M. (2007). The use of radio frequency ID tags to track ore in mining operations. In Proceedings: Application of Computers and Operations Research in Mineral Industry, Santiago, Chile, 601–606

Rossi, M.E., and Deutsch, C. V (2014). Mineral Resource Estimation. Springer, Dordrecht

Ruder, S. (2016). An overview of gradient descent optimization algorithms. ArXiv, 1609.04747, 1–14

Sarma, P., Durlofsky, L.J., Aziz, K., and Chen, W.H. (2006). Efficient real-time reservoir management using adjointbased optimal control and model updating. Computational Geosciences, 10(1), 3–36

Shirangi, M.G. (2017). Advanced techniques for closed-loop reservoir optimization under uncertainty, Ph.D. Thesis. Standford University

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484–489

Vargas-Guzmán, J.A., and Dimitrakopoulos, R. (2002). Conditional simulation of random fields by successive residuals. Mathematical Geology, 34(5), 597–611

Vo, H.X., and Durlofsky, L.J. (2014). A new differentiable parameterization based on principal component analysis for the low-dimensional representation of complex geological models. Mathematical Geosciences, 46(7), 775–813

Xue, L., and Zhang, D. (2014). A multimodel data assimilation framework via the ensemble Kalman filter. Water Resources Research, 50(5), 4197–4219

Yüksel, C., Benndorf, J., Lindig, M., and Lohsträter, O. (2017). Updating the coal quality parameters in multiple production benches based on combined material measurement: a full case study. International Journal of Coal Science and Technology, 4(2), 159–171