



# Adaptive self-learning mechanisms for updating short-term production decisions in an industrial mining complex

Ashish Kumar<sup>1</sup> · Roussos Dimitrakopoulos<sup>1</sup> · Marco Maulen<sup>2</sup>

Received: 26 March 2019 / Accepted: 16 March 2020 / Published online: 2 April 2020  
© The Author(s) 2020

## Abstract

A mining complex is an integrated value chain where the materials extracted from a group of mineral deposits are sent to different processing streams to produce sellable products. A major short-term decision in a mining complex is to determine the flow of materials that first includes deciding which handling facilities to send the extracted materials and then determining how to utilize the processing facilities. The flow of materials through the mining complex is significantly dependent on the performance of and interaction between its different components. 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. This paper proposes a new continuous updating framework that combines policy gradient reinforcement learning and an extended ensemble Kalman filter to adapt the short-term flow of materials in a mining complex with incoming information. The framework first uses a new extended ensemble Kalman filter to update the uncertainty models of the different components of a mining complex with new incoming information. Then, the updated uncertainty models are fed to a neural network trained using a policy gradient reinforcement learning algorithm to adapt the short-term flow of materials in a mining complex. The proposed framework is applied to a copper mining complex and shows its ability to efficiently adapt the short-term flow of materials in an operational mining environment with new incoming information. The framework better meets the different production targets while improving the cumulative cash flow compared to industry standard approaches.

**Keywords** Mining complex · Production planning · Artificial intelligence · Reinforcement learning · Sensor information · Ensemble Kalman filter · Real-time · Destination policies · Deep learning

## Introduction

A mining complex is an integrated value chain network with multiple interlinked components including suppliers of raw

materials (mineral deposits and external inventories), heavy earth moving equipment (shovel, trucks, and conveyor belts), handling facilities (crushers, stockpiles, and waste dumps), processing facilities (mineral processing mills and leach pads), and customer/commodity markets. Uncertainty is a characteristic of a mining complex, starting from the supply of different types of raw materials extracted from the mineral deposits involved (Dimitrakopoulos et al. 2002). Stochastic optimization models account for uncertainty and generate production decisions that yield higher value and manage the technical risk of not meeting the production targets (Mai et al. 2019; Matamoros and Dimitrakopoulos 2016). 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, 2018). The short-term production plan determines the daily/weekly/monthly production

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10845-020-01562-5>) contains supplementary material, which is available to authorized users.

---

✉ Ashish Kumar  
ashish.kumar@mail.mcgill.ca

Roussos Dimitrakopoulos  
roussos.dimitrakopoulos@mcgill.ca

Marco Maulen  
Marco.Maulen1@bhpbilliton.com

<sup>1</sup> COSMO – Stochastic Mine Planning Laboratory, Department of Mining and Materials Engineering, McGill University, FDA Building, 3450 University Street, Montreal, QC H3A 0E8, Canada

<sup>2</sup> Practice Lead Mining Technical, BHP, Santiago, Chile

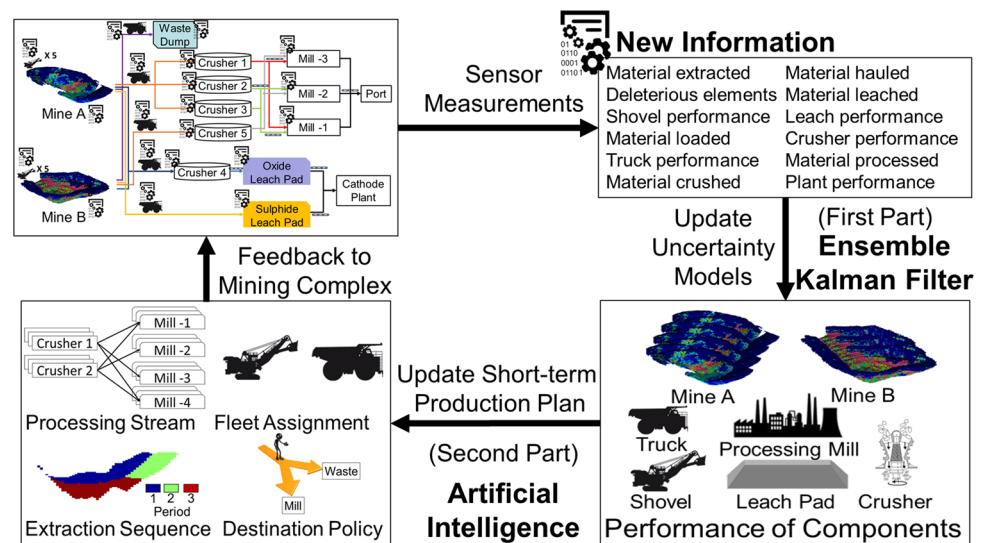
decisions within the long-term production plan to meet annual targets. A review of short-term production planning in mining operations can be found in Blom et al. (2019). In addition to supply uncertainty, the short-term production plan accounts for uncertainty in the performance of equipment to determine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

New digital technologies, including the development of advanced sensors and monitoring devices, have enabled the acquisition of new information about the performance of the different components of a mining complex that affect the flow of materials in a value chain. Sensors installed on drills, shovels, trucks, conveyor belts, crushers, and mineral processing mills (Dalm et al. 2014, 2018; Goetz et al. 2009; Iyakwari et al. 2016; Wambeke and Benndorf 2018) continuously measure the performance of the mining equipment and processing streams (processing and handling facilities), as well as different pertinent properties of the materials being handled. In addition to the new sensor information, conventional sources of new information include blasthole sampling that determines the pertinent properties of materials extracted (Rossi and Deutsch 2013), monitoring devices that measure the performance of equipment (Koellner et al. 2004), and tracking devices that track the location of materials (Brewer et al. 1999; Rosa et al. 2007).

The core existing technologies can only integrate new information that is conventionally collected, such as grade control that integrates blasthole data to identify ore/waste boundaries in the blasted areas of mineral deposits (Dimitrakopoulos and Godoy 2014; Verly 2005) or dispatching stations that monitor the equipment for assignment and dispatch decisions (Kargupta et al. 2010; Nguyen and Bui 2015). However, these technologies are unable to integrate the sensor-generated information to adapt the short-term production plan. A continuous updating framework, shown in Fig. 1, is needed to adapt the short-term production plan of a mining complex with new information generated from both sensors and conventional sources. The continuous updating framework consists of two parts. First, the new information generated from the different sources in a mining complex is used to update the performance of its different components, which includes uncertainty in the supply of materials

from the mineral deposits, the performance of equipment, and the processing streams' capabilities (productivity, recovery, etc.). Second, the updated performance of the different components of a mining complex is then fed to an artificial intelligence framework, which, in the present work, is a neural network agent that is trained using policy gradient reinforcement learning to adapt the short-term production plan. The adapted short-term production plan is fed back to the mining complex to generate updated production forecasts. The adapted production plan is then followed, more sensor data is collected as the mining operations progress, and the production plan is adapted again, and the cycle continues. Benndorf and Buxton (2016) proposed a framework to update the mine planning decisions with new information. Related is also the work of Hou et al. (2015) and Shirangi (2017), who proposed a continuous updating framework to update the production plan of smart oil fields. However, the existing frameworks, both in mine planning and smart oil fields, require re-optimization of the production plan, which is computationally expensive with the available optimization techniques. Lamghari (2017) provided a detailed review of the different techniques used for production planning in mining complexes and smart oil fields. The new information generated in a mining complex can be categorized as "soft" and "hard" data, based on the precision of their measurement. Sensor-generated information is "soft" data because it is noisy, uncertain, and ambiguous when collected during operations from different components of a mining complex. Direct measurements, such as those derived from drillhole samples, which are analyzed in geochemical laboratories and are substantially more precise, are considered "hard" data. Consequently, the first part of the continuous updating framework in a mining complex, as shown in Fig. 1, aims at generating updated uncertainty models of the different components of a mining complex that are consistent with the hard data and minimize the mismatch between (a) the observed and forecasted production data, as well as (b) the soft and hard data. Evensen et al. (1994) proposed the ensemble Kalman filter (EnKF) that updates the non-linear processes with new information and has long been used for petroleum reservoir flow simulation and production forecasting (Dovera and Della Rossa 2011; Kumar and Srinivasan 2019; Xu and Hernández 2019; 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; Dalm et al. 2018; Yüksel et al. 2018). Methods such as randomized maximum likelihood (Chen and Oliver 2012; Sarma et al. 2006; Shirangi 2017; Vo and Durlafsky 2014) and Markov mesh models (Panzeri et al. 2016) are also used to update the pertinent

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



petroleum reservoir-related attributes. Vargas-Guzmán and Dimitrakopoulos (2002) and Jewbali and Dimitrakopoulos (2011) proposed a column-wise decomposition of the covariance matrix (CSSR) to update the pertinent attributes of mineral deposits with new hard data. However, the CSSR method cannot integrate the soft information generated from sensors. The outlined methods for updating pertinent attributes of mineral deposits with EnKF and CSSR are limited to a single attribute. This paper presents a new extension of EnKF that allows the updating of multiple correlated attributes in mineral deposits with minimum/maximum autocorrelation factors (Desbarats and Dimitrakopoulos 2000).

The second part of the updating framework (Fig. 1) aims at adapting the short-term production plan of a mining complex with the updated uncertainty models of its different components. Reinforcement learning methods are efficient in decision-making with new information. In recent years, reinforcement learning-based methods have shown exceptional performance at generating neural network agents that are capable of making very efficient decisions for different complex environments (Aissani et al. 2012; Barde et al. 2019; Mnih et al. 2013; Silver et al. 2016). Paduraru and Dimitrakopoulos (2018) proposed a Bayesian reinforcement learning algorithm to optimize the destination policies of materials in a mining complex. However, the method developed requires a predefined extraction sequence to calculate the expected a posteriori improvement in the objective function during the optimization. Paduraru and Dimitrakopoulos (2019) proposed a policy gradient reinforcement learning algorithm to optimize the neural network destination policies of materials in a mining complex while accounting for supply and equipment performance uncertainty. The neural network destination policies increased the expected NPV by 6.5% compared to the mine's cut-off grade destination policies for a copper mining complex. However, the method is (a)

limited to a single product mining complex, and (b) does not provide a required continuous updating of the short-term production plan regarding destination policies of materials with the new information generated from sensors and/or conventional sources.

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

## Methods

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

### Updating stochastic orebody simulations

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

### Updating algorithm

A mineral deposit is discretized into an array of three-dimensional volumes referred to as mining blocks. The mining blocks are further discretized into multiple internal nodes. Let  $\mathbb{Z}_e^{t',s}(x)$  be a realization  $s \in S$  of the vector of the spatial random field consisting of elements  $\mathbb{Z}_e^{t',s}(x_i)$ .  $\mathbb{Z}_e^{t',s}(x_i)$  represents the simulated MAF value of element  $e$  at location  $x_i$ , at time  $t'$ , under scenarios  $s$ , with  $i \in [1, \mathcal{N}]$ , being the index of internal nodes. Initial ensembles of MAF values are represented by  $\mathbb{Z}_e^{t',s}(x)$  for the multiple elements in the mineral deposit. Let matrix  $A_{t'}$  describe the contribution of each internal node at the location  $x_i$  at time  $t'$ , towards the new information observed in the mining complex. The new information observed at the time  $t'$  is also de-correlated using MAF into MAF factor  $l_e^{t'}$  for element  $e$ . The Gaussian assumption in the ensemble Kalman filter is handled by transforming  $\mathbb{Z}_e^{t',s}(x)$ ,  $\mathbb{Z}_e^{t',s}(x_i)$ , and  $l_e^{t'}$  using the Gaussian anamorphosis function  $\Phi_G^e$ . The transformed vectors,  $U_e^{t',s}(x)$ ,  $u_e^{t',s}(x_i) = \Phi_G^e(\mathbb{Z}_e^{t',s}(x_i))$ , and  $m_e^{t'} = \Phi_G^e(l_e^{t'})$  are then

used in the EnKF updating process.  $U_e^{t',s}(x)$  is the vector of elements  $u_e^{t',s}(x_i)$ . A random noise  $\epsilon_e^{t'}$  is added in the new information to represent the noise with the measurement of new information (Eq. 1). The model-based prediction  $P_e^{t',s}$ , which represents the predictions based on initial ensembles at the location of observed information is given by Eq. 2.

$$o_e^{t'} = m_e^{t'} + \epsilon_e^{t'}, \quad \forall e \in E \quad (1)$$

$$P_e^{t',s} = A_{t'} \cdot U_e^{t',s}(x), \quad \forall e \in E, \quad s \in S \quad (2)$$

$$U_e^{t'+1,s}(x) = U_e^{t',s}(x) + K_e^{t'} \cdot (o_e^{t'} - P_e^{t',s}), \quad \forall e \in E, \quad s \in S \quad (3)$$

$$K_e^{t'} = \left( A_{t'}^T \cdot C_{u_e u_e}^{t'} \cdot A_{t'} + C_{o_e o_e}^{t'} \right)^{-1} A_{t'}^T \cdot C_{u_e u_e}^{t'}, \quad \forall e \in E \quad (4)$$

EnKF uses Eq. 3 to update the initial ensembles ( $U_e^{t',s}(x)$ ) with the new information based on the difference between new information and model-based prediction, and the Kalman gain. The Kalman gain  $K_e^{t'}$  is calculated using Eq. 4 and defines the significance of the model compared to the new information through the error covariance matrix of the model ( $C_{u_e u_e}^{t'}$ ) and observations ( $C_{o_e o_e}^{t'}$ ). For instance, if the new information is inaccurate, then the term  $C_{o_e o_e}^{t'}$ , will be high, which results in low Kalman gain.

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

$$C_{u_e u_e}^{t'}(x_i) \cong \frac{1}{S} \sum_{s=1}^S \left( u_e^{t',s}(x_i) - \overline{u_e^{t',s}(x_i)} \right) \cdot \left( \left( u_e^{t',s}(x_i) - \overline{u_e^{t',s}(x_i)} \right) \right)^T, \quad \forall i \in \mathcal{N}, \quad e \in E \quad (5)$$

EnKF approximates the model error covariance matrix with a finite set of ensembles (Eq. 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  $\Phi_G^{e-1}(U_e^{t'+1,s}(x))$  to generate updated MAF ensemble values  $\mathbb{Z}_e^{t'+1,s}(x)$ . The updated MAF ensemble values are further back-transformed using the MAF inverse transformation

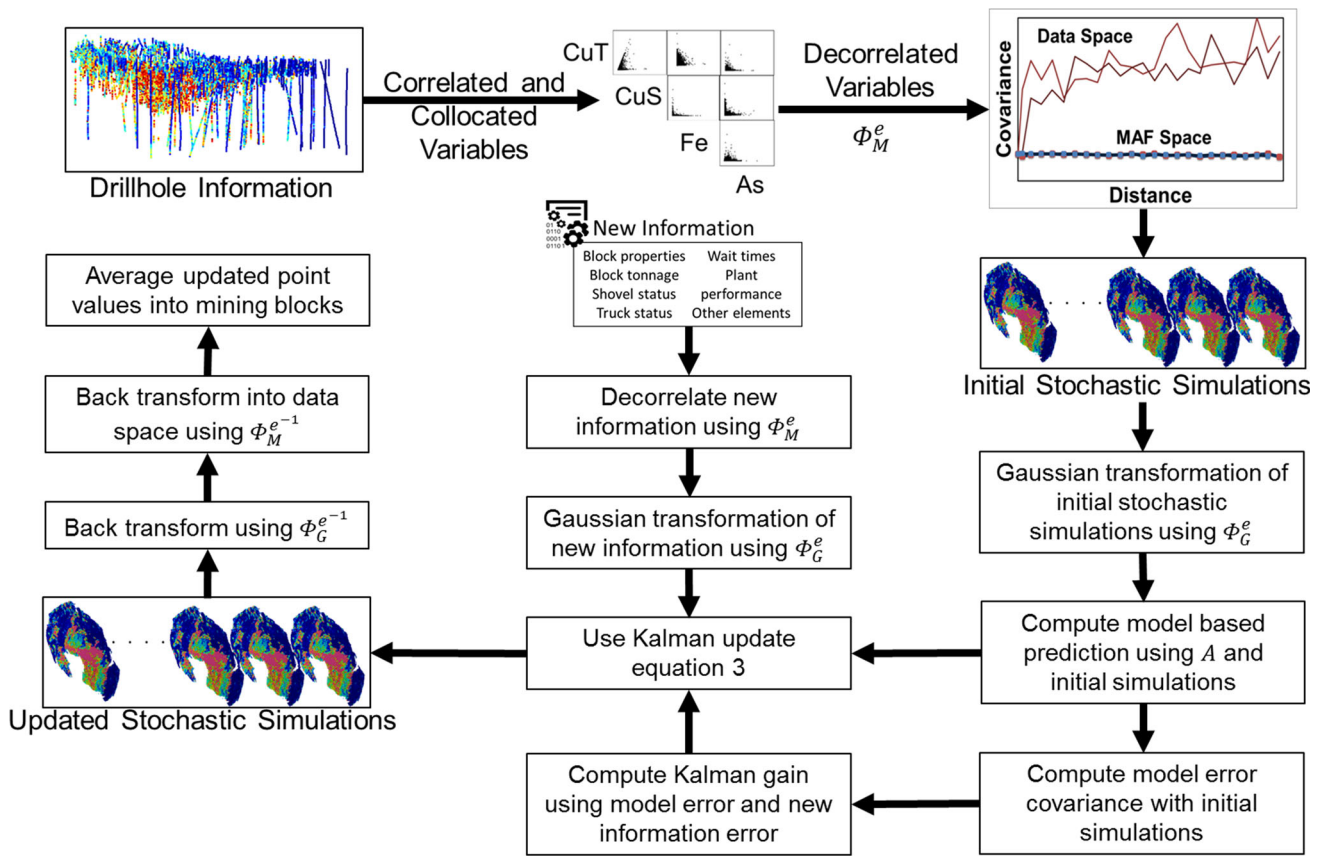


Fig. 2 Updating stochastic simulations of mineral deposits with new information

function and averaged to generate values of different elements in the mining blocks for different ensembles (Eq. 6).

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

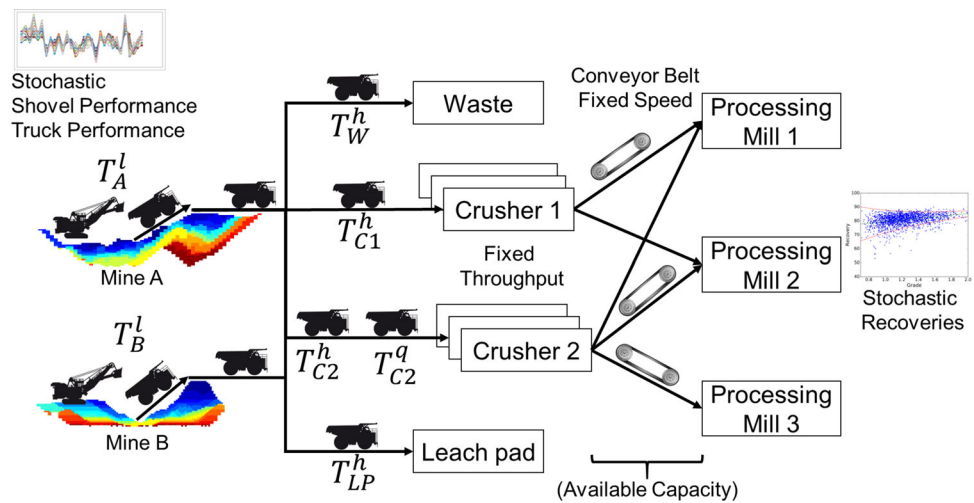
### Updating short-term destination policies in a mining complex

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

### A stochastic model of a mining complex

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

**Fig. 3** An illustrative example of a stochastic model of a mining complex



ing time to a destination  $T_{m,d,s}^h$ , and wait time at a destination  $T_{d,s}^q$ , and is calculated using Eq. 7.

$$T_{m,d,s}^E = f \left( T_{m,s}^l, T_{m,d,s}^h, T_{d,s}^q \right), \quad \forall m \in \mathcal{M}, \quad d \in \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{W}, \quad s \in \mathbb{S} \quad (7)$$

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

**Updating algorithm**

The stochastic model of a mining complex presented in “A stochastic model of a mining complex” section simulates the flow of materials in the mining complex under the joint uncertainty scenario  $\mathbb{S}$ , which is used to train the neural network destination policies. Note, the proposed model decides the destination of materials based on multiple elements in a mining complex, given a fixed extraction sequence. The complete training process of the neural network is presented in Fig. 4a. The joint uncertainty scenarios are fed to the stochastic model to perform the extraction and hauling simulations that generate information about the input state ( $SV_i$ ), which includes the quality and quantity of materials extracted, hauled, crushed, leached, and discarded under joint uncertainty scenarios.  $SV_i$  is fed to input neurons in the fully connected feed-forward neural network. The input to different hidden neurons ( $h_j$ )

is calculated using Eq. 8. Equation 9 is used to calculate the output of hidden neurons using the rectified linear function (Nair and Hinton 2010). The input to output neurons ( $o_k$ ) is then calculated using Eq. 10. The weight matrix  $w_{ij}^h$  and  $w_{jk}^o$  represent the weight associated with arcs from input ( $i$ ) to hidden ( $j$ ) and hidden to output ( $k$ ) neurons.

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

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

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

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

The output from output neurons defines the decisions variable  $z_{b,d,t}$ , that determines if (1) or not (0) a block  $b$  is sent to a destination  $d$  in a period  $t$  and is calculated using Eq. 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 property  $a$ , and mass respectively at the different destinations  $i$ .

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

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

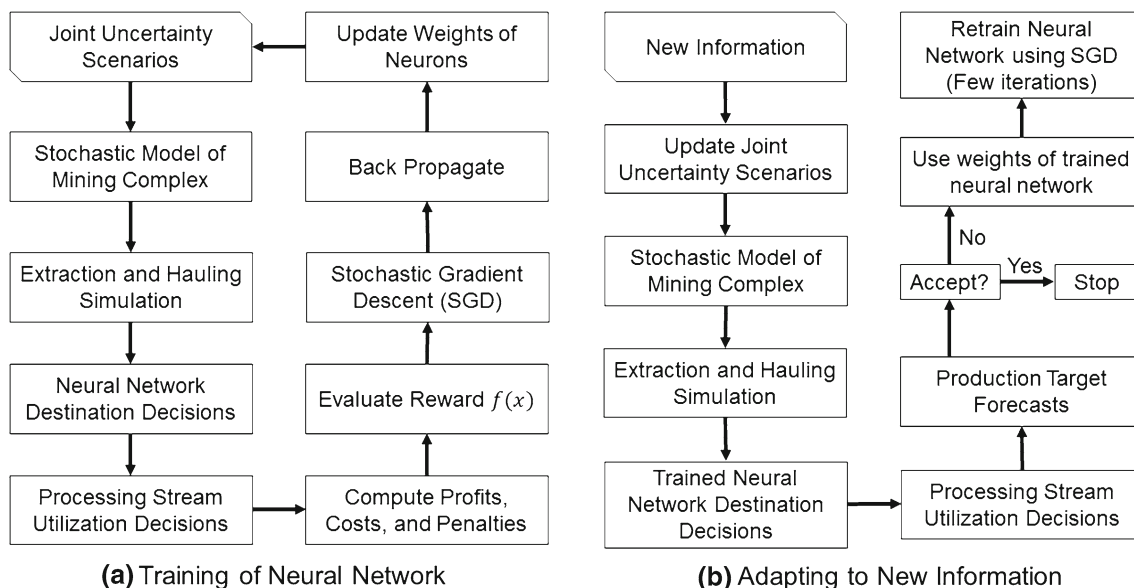


Fig. 4 Process of training the neural network and adapting to new information

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

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

$$\sum_{j \in \mathcal{P} \cup \mathcal{L}_O} y_{a,i,j,t,s} = 1, \quad \forall t \in \mathbb{T}, \quad i \in \mathcal{C}, \quad s \in \mathbb{S} \tag{15}$$

$$v_{a,i,t,s} - d_{a,i,t,s}^+ \leq U_{a,i,t}, \quad \forall t \in \mathbb{T}, \quad a \in \mathbb{P}_M, \quad i \in \mathcal{P} \cup \mathcal{L}_S \cup \mathcal{L}_O, \quad s \in \mathbb{S} \tag{16}$$

$$v_{a,i,t,s} + d_{a,i,t,s}^- \geq L_{a,i,t}, \quad \forall t \in \mathbb{T}, \quad a \in \mathbb{P}_M, \quad i \in \mathcal{P} \cup \mathcal{L}_S \cup \mathcal{L}_O, \quad s \in \mathbb{S} \tag{17}$$

$$f(X) = \underbrace{\frac{1}{\mathbb{S}} \sum_{s \in \mathbb{S}} \sum_{t \in \mathbb{T}} \sum_{i \in \mathcal{P} \cup \mathcal{L}_O \cup \mathcal{L}_S} \sum_{a \in \mathbb{P}_M} P_{a,i} \cdot v_{a,i,t,s} \cdot r_{a,i,s}}_{\text{Part I}} - \underbrace{\frac{1}{\mathbb{S}} \sum_{s \in \mathbb{S}} \sum_{t \in \mathbb{T}} \sum_{i \in \mathcal{P} \cup \mathcal{L}_S \cup \mathcal{L}_O} \sum_{a \in \mathbb{P}_T} C_{a,i} \cdot v_{a,i,t,s}}_{\text{Part II}} - \underbrace{\frac{1}{\mathbb{S}} \sum_{s \in \mathbb{S}} \sum_{t \in \mathbb{T}} \sum_{i \in \mathcal{P} \cup \mathcal{L}_S \cup \mathcal{L}_O} \sum_{a \in \mathbb{P}_M} (c_{a,i}^+ \cdot d_{a,i,t,s}^+ + c_{a,i}^- \cdot d_{a,i,t,s}^-)}_{\text{Part III}} \tag{18}$$

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/cash flow/reward function is given by Eq. 18. Part I in the objective function represents the profits from selling different products; Part II represents the different costs incurred throughout the flow of materials, and Part III represents the penalties incurred due to deviation from different production targets. The objective function is an expected value. Equations 12–18 are based on recent developments in stochastic mine planning models (Goodfellow and Dimitrakopoulos 2016; Montiel and Dimitrakopoulos 2015; Quigley and Dimitrakopoulos 2019). Policy gradient reinforcement learning (Sutton et al. 2000) offers the ability that, given a reward function  $f$  and probability density function  $z_W$  parameterized by  $W$ , the equality in Eq. 19 below holds true.

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

$f(x)$  in Eq. 19 corresponds to the reward function and  $z_W(x)$ , corresponds to the action-selection probabilities computed using Eq. 11. The weight matrix  $W$  contains the values of the hidden  $w_{ij}^h$ , and the output neurons  $w_{jk}^o$ . As it is common in stochastic gradient methods (Bottou 2010),  $E_{x \sim z_W(x)}[f(x) \nabla_W \log(z_W(x))]$  is replaced with  $f(X) \nabla_W \log(z_W(X))$ , where  $f(X)$  represent the cumulative reward obtained during the planning horizon  $\mathbb{T}$  using the vector of decisions  $X$ . The gradient of  $\log(z_W(X))$  can, therefore, be calculated using Eq. 20, where the sum is over the planning horizon and over the destinations. Finally, the stochastic approximation of  $\nabla_W E_{x \sim z_W(x)}[f(x)]$  can be computed using Eqs. 18–20.

$$\nabla_W \log(z_W(X)) = \sum_{t \in \mathbb{T}} \sum_{d \in \mathcal{C} \cup \mathcal{L}_S \cup \mathcal{W}} \nabla_W \log z_W(d) z_{b,d,t} \quad (20)$$

$$g_{i+1} = \gamma g_i + (1 - \gamma) \nabla_W E_{x \sim z_W(x)}[f(x)]^2, \quad \forall i \in [1, nIter] \quad (21)$$

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

The weight matrix  $W = \{w_{ij}^h, w_{jk}^o\}$  of the neurons in the neural network is initialized randomly and updated using the gradient ascent method named RMSprop (Hinton et al. 2012). The RMSprop method uses Eqs. 21 and 22 to backpropagate and update the weight of the neurons in the training phase of the neural network. This process (Eqs. 8–22) continues, and the neural network is trained until the pre-defined stopping criteria ( $nIter$ ) are reached.

The training phase of the neural network allows the generation of destination policies that can adapt to new information. Figure 4b 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 “A stochastic model of a mining complex” section. The updated joint uncertainty scenarios are then fed to the stochastic model outlined in “A stochastic model of a mining complex” section, which simulates the extraction and hauling of materials. The information from the previous step is fed to the trained neural network that decides the destination of materials and the materials from such destinations are then sent to one of the processing streams based on the available capacity of the different processing streams. Finally, the forecasts for the different production targets are calculated using Eqs. 12–18 and further evaluated regarding their probability of meeting the different production targets.

The neural network is retrained for a few iterations if the production targets are not met to adjust the weight of the neural network and better meet the production targets.

## Application at a copper mining complex

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

## Overview of the copper mining complex

The copper mining complex consists of two mineral deposits (A and B) with mining blocks of size  $25 \times 25 \times 15 \text{ m}^3$ . The mineralization has eight different mine zones each. The materials are extracted from both deposits and are sent to one of the seven destinations (five crushers, one sulphide leach pad, and one waste dump), as shown in Fig. 5. For measuring the performance of the proposed framework, a part of the deposit that consists of 5581 mining blocks in each deposit extracted over 210 days is used. Materials from five different crushers are then processed at three different processing mills and an oxide leach pad.

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



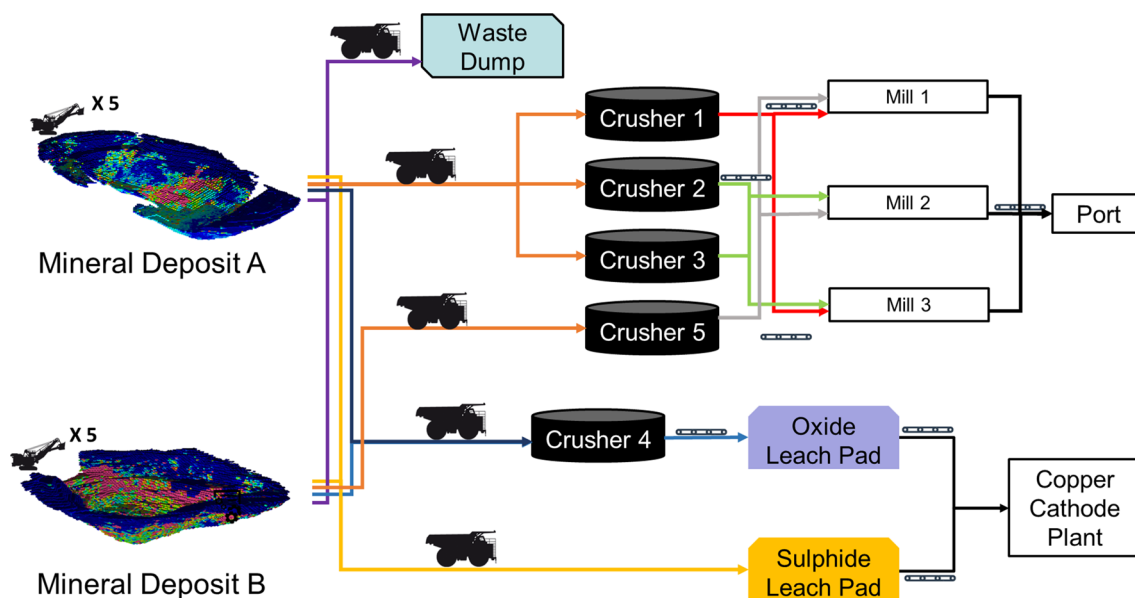


Fig. 5 The copper mining complex

### Cut-off grade versus adaptive neural network destination policies

The copper mining complex currently uses a single element (copper) predefined cut-off grade based destination policies optimized using Lane’s theory (Lane 1984, 1988; Rendu 2014) and cannot account for new information collected during the mine’s operation. The copper mining complex is a major producer of copper products and does not consider secondary products in the optimization of its cut-off grade destination policies. The details of the cut-off grade destination policies are outlined in Table 1. First, the materials are classified as sulphide high grade (SHG), sulphide low grade (SLG), oxide based on the materials classification criteria [i.e., ratio of soluble copper (CuS) to total copper (CuT)]. The materials classification criteria are necessary to determine the possible processing destinations allowed to process the materials. The cut-off grade destination policies then use the cut-off grades specified in Table 1 to determine the destination at which the material will be processed.

The neural network destination policies decide the destination of mining blocks based on the properties of multiple elements in a mining block, as well as the performance of and interaction between the different components of the mining complex. In addition, the proposed method adapts such destination decisions of mining blocks with new incoming information in the mining complex (see “Updating algorithm” section). Similar to the cut-off grade destination policies, the materials are first characterized as SHG, SLG, oxide, and waste, based on the material classification criteria mentioned in Table 1 to find the allowed processing destinations for a mining block. However, instead of using the

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

Materials classification	Materials classification criteria	Cut-off grade destination policies	Destination
SHG	$\frac{CuS}{CuT} \leq 0.2$	$CuT \geq 0.6$	Processing mill
		$0.3 \leq CuT < 0.6$	Sulphide leach pad
		$CuT < 0.3$	Waste dump
SLG	$0.2 < \frac{CuS}{CuT} < 0.5$	$CuT > 0.3$	Sulphide leach pad
		$CuT \leq 0.3$	Waste dump
Oxide	$\frac{CuS}{CuT} \geq 0.5$	$CuS \geq 0.2$	Oxide leach pad
		$CuS < 0.2$	Waste dump

cut-off grade destination policies mentioned in Table 1, the neural network destination policies are used to decide the destination of such materials. Three different neural networks are built and trained using policy gradient reinforcement learning. As mentioned in “Updating short-term destination policies in a mining complex” section, the neural networks decide whether (1) or not (0) to process the materials at (1) the processing mills, (2) a sulphide leach pad, or (3) an oxide leach pad.

### Parameter selection

This section discusses the selection of different parameters associated with the proposed adaptive neural network desti-

nation policies. The state vector information  $SV_i$  consists of 7–32 different types of information depending on the complexity of the processing destination and are fed to the input neurons of the neural network. For instance,  $SV_i$  for the processing mill, neural network consists of information about the mass of a mining block, different elements such as total copper, soluble copper, arsenic, gold, silver, and molybdenum in the mining block, the materials being crushed and leached, the performance of equipment, and the wait times at the crushers. Similarly, the number of hidden neurons in the neural network ranges from 300 to 800, depending on the number of input neurons. There are only two output neurons to decide whether (1) or not (0) the mining block is processed at the respective destination. The learning rate and the decay rate with the neural network is set to  $10^{-3}$  and 0.99, respectively, as suggested in Hinton et al. (2012). The smoothing term is set to  $10^{-6}$  (Ruder 2016). The weight of the neurons in the neural network is initialized randomly using the Xavier initialization (Glorot and Bengio 2010). The number of iterations required to train the neural network is set to 7500. The number of mineral deposit simulations to use for training the neural network is set to 15 based on the tests show in the supplementary material. The details of parameter selection are also presented in the supplementary material of the present manuscript.

## Results

The results of the proposed adaptive neural network destination policies to update the short-term destination decisions with new information are presented in this section. Results are reported using the 10th, 50th, and 90th percentile risk profiles (P10, P50, and P90 respectively) of the different performance indicators considering 100 joint uncertainty scenarios (10 equipment performance and 10 orebody scenarios). The results reported in this section are based on a set of 100 joint uncertainty scenarios that were not used to train the neural network destination policies. Testing the neural network destination policies on an unseen set of joint uncertainty scenarios shows the reliability of the proposed framework and highlights the overfitting issues, if any, with the neural network destination policies. The forecasts of the production targets with the proposed framework are compared to the forecasts of the cut-off grade destination policies over the same 100 joint uncertainty scenarios throughout its presentation and discussion to highlight the differences and added value of the adaptive framework, where appropriate. The training phase of the neural network takes about 52 h, with 12,500 iterations on an Intel processor core i7 with 8 GB of RAM. However, it only takes about 5 min to update the stochastic simulations of the two mineral deposits and to adapt the destination decisions of mining blocks for 210 days using the proposed adaptive framework. The results are pre-

sented for both the destination policies for initial and update stochastic simulations of mineral deposits. The results presented for metal production and cash flows are scaled for confidentiality purposes (mine's cut-off grade based destination policies for initial simulations being 100%). Additional results from the case study are presented in the supplementary material.

## Updated stochastic simulations of mineral deposit

Figure 6 shows one of the initial and updated simulations of the total copper mineral attribute of the mineral deposit A at block support. The initial stochastic simulations of six correlated elements in the two mineral deposits, conditional to the exploration drillholes' samples, are generated using a generalized sequential Gaussian simulation (Dimitrakopoulos and Luo 2004). Six different correlated elements: soluble copper, total copper, arsenic, gold, silver, and molybdenum, in the two mineral deposits are updated using the method discussed in "Updating stochastic orebody simulations" section with the new blasthole data collected during the short-term operations. The blasthole data in a mine zone are only considered to update the mining blocks in the same mine zone to respect the geological features of the mineral deposit. It is clear from Fig. 6 that the updated simulations maintain the significant structures inferred from the exploration drillholes 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 mineral deposit A is very different for the initial and updated simulations. The updated simulations show a higher proportion of high-grade copper materials, as compared to the initial simulations.

## Production targets

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

Figure 7b shows the risk profile of meeting the capacity target with mill-2 for initial simulations using neural network destination policies compared to the cut-off grade destination policies in Fig. 7a. The neural network destination policies are better at meeting the target with maximum utilization of the mill's capacity, as compared to high fluctuations and lower chances of meeting the target in the cut-off grade destination policies. The neural network destination policies (Fig. 7d) has increased the chance of meeting production targets compared to the high fluctuations in the cut-off grade destination policies (Fig. 7c) over the updated simulations. Figure 8a, b show the risk of meeting the blending target of arsenic at mill-2 for initial simulations with neural network and cut-off grade destination policies, respectively.

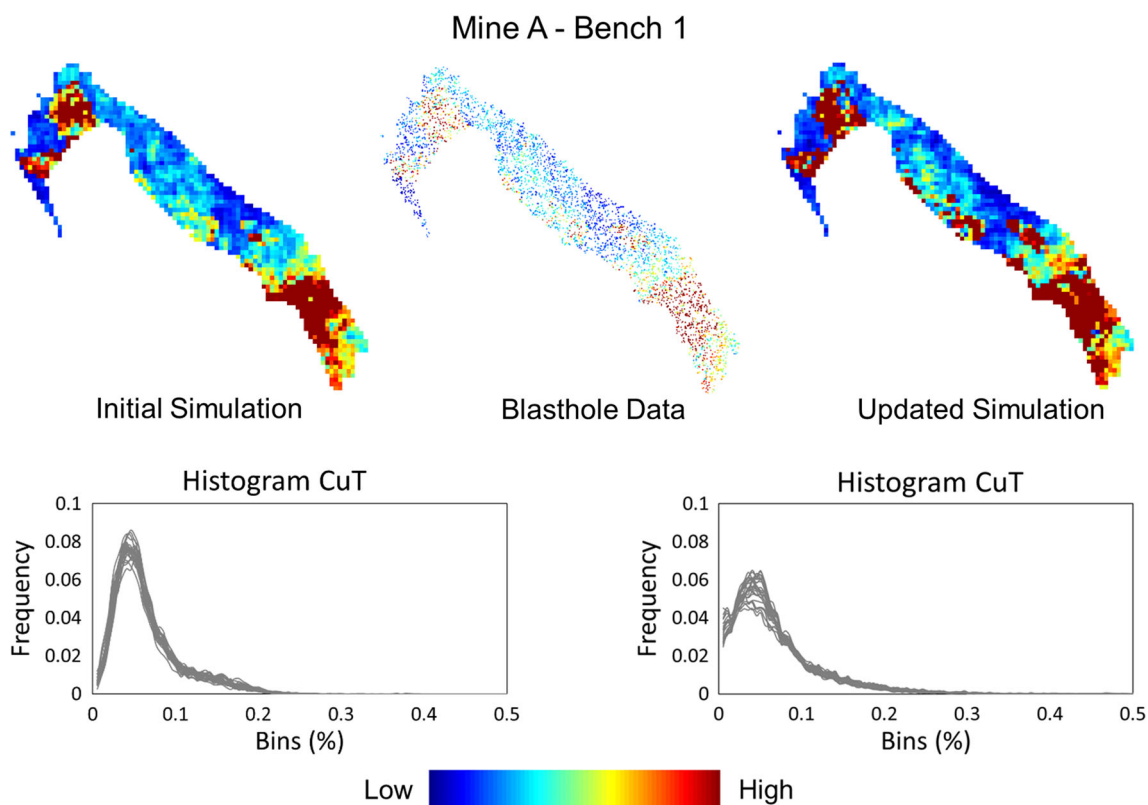
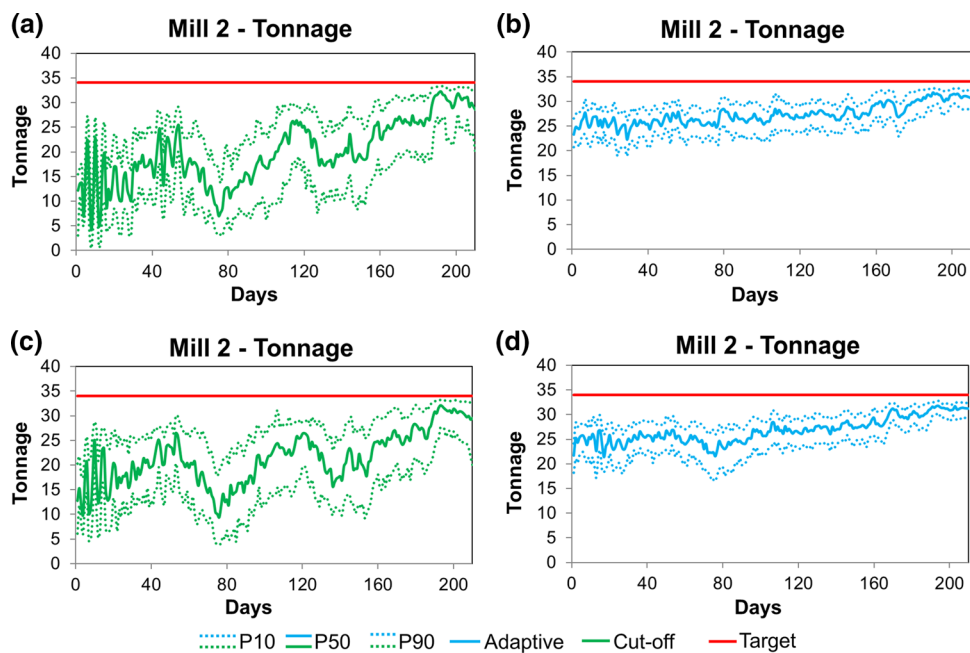


Fig. 6 Updated block simulations compared to initial block simulations for bench 1 for the mineral deposit A

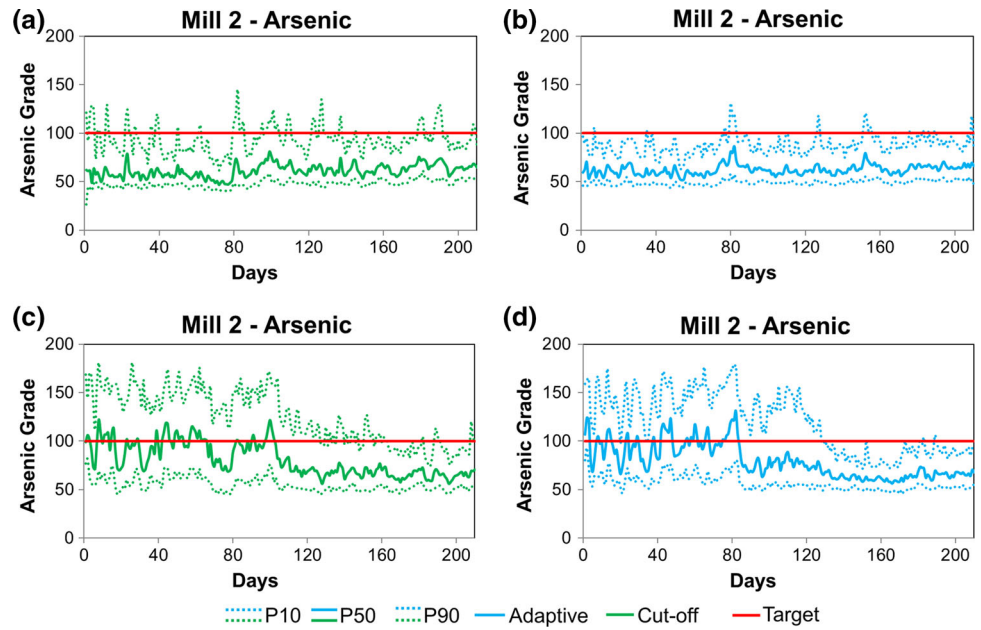
Fig. 7 Forecasts of the capacity target of mill-2 with the a initial cut-off grade block destinations, b initial neural network block destinations, c updated cut-off grade block destinations, and d updated neural network block destinations



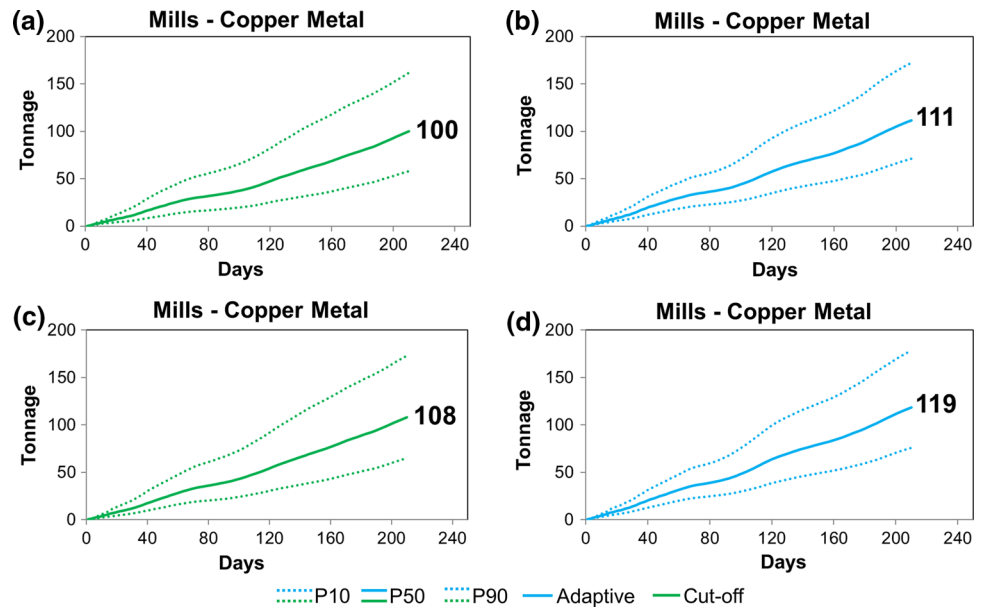
The neural network destination policies have higher chances of meeting such a target with minimal deviations only after 80 days, as compared to the cut-off grade destination policies, which have a higher chance of deviating from such targets, more specifically during the first 80 days. The two destina-

tion policies are unable to meet the blending restrictions as shown in Fig. 8c, d over the updated simulations. The lower chances of meeting the arsenic target with the updated destination decisions are due to the fixed extraction sequence decision in the proposed framework. Therefore, if there is a

**Fig. 8** Forecasts of arsenic blending target of mill-2 with the **a** initial cut-off grade block destinations, **b** initial neural network block destinations, **c** updated cut-off grade block destinations, and **d** updated neural network block destinations



**Fig. 9** Forecasts of total copper production at the processing mills with the **a** initial cut-off grade block destinations, **b** initial neural network block destinations, **c** updated cut-off grade block destinations, and **d** updated neural network block destinations



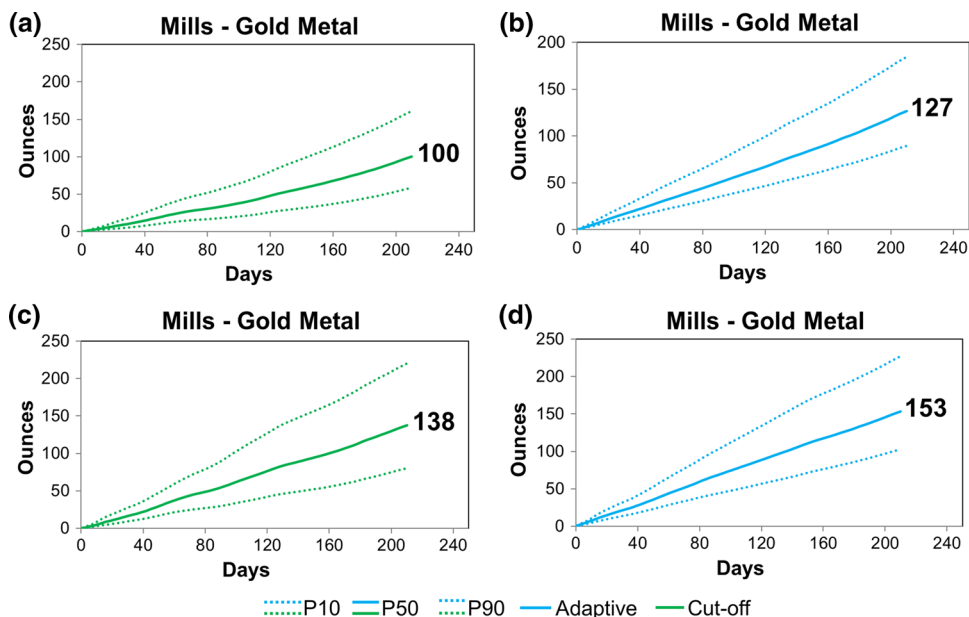
high concentration of arsenic in the updated simulations, it is hard to control the arsenic concentration in the mill without adapting the extraction sequence.

**Metal production**

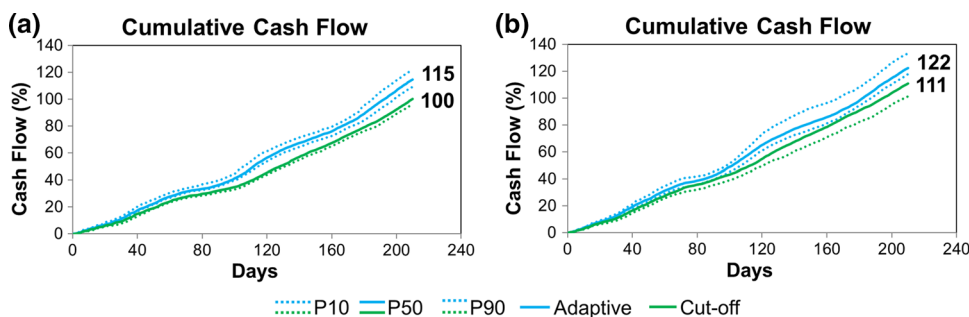
Figure 9a, b represent the risk profile of cumulative copper production at the mills for the initial simulations with neural network and cut-off grade destination policies, respectively. The neural network destination policies recover 11% additional copper metal, as compared to the mine’s cut-off grade destination policies for the initial simulations. The neural network destination policies recover an additional 19% cop-

per metal (Fig. 9d), as compared to an additional 8% copper metal in the mine’s cut-off grade destination policies (Fig. 9c) over the updated simulations. Figure 10 shows the risk profiles of the production of secondary product gold concentrate using the neural network and the cut-off grade destination policies. The neural network destination policies generate 27% additional gold product (Fig. 10b), as compared to the mine’s cut-off grade destination policies (Fig. 10a) over the initial simulations. The adapted decisions of neural network destination policies generate an additional 53% of the gold product (Fig. 10d), as compared to an additional 38% for the mine’s cut-off grade destination policies (Fig. 10c) over the updated simulations.

**Fig. 10** Forecasts of total gold production at the processing mills with the **a** initial cut-off grade block destinations, **b** initial neural network block destinations, **c** updated cut-off grade block destinations, and **d** updated neural network block destinations



**Fig. 11** Forecasts of the cumulative cash flow of the mining complex with the **a** initial cut-off grade and neural network block destinations, and **b** updated cut-off grade and neural network block destinations



**Cash flows**

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

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

**Updated destination decisions**

Figure 12b shows the destination decisions of the neural network destination policies compared to the cut-off grade destination policies in Fig. 12a for initial simulations. The adapted destination decisions of the neural network and the cut-off grade destination policies are shown in Fig. 12c, d, respectively. The neural network destination decisions are very different from the cut-off grade destination decisions for initial and update simulations, which result in better chances

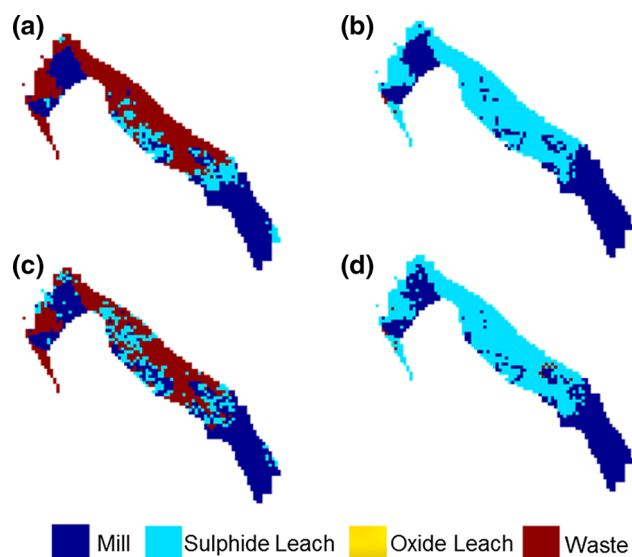
of meeting production targets, consistently higher cumulative cash flows, and increased metal production.

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

1. 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. Integrate multiple sources of uncertainty, such as the supply of materials, the performance of equipment, and the recovery of metal during the decision-making process
3. Account for multiple products, such as copper, gold, silver, and molybdenum, as well as deleterious elements such as arsenic, while deciding the destination of mining blocks.

**Conclusions**

This paper presents a novel continuous updating framework for adapting the short-term flow of materials in a mining complex with new incoming information. The framework



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

consists of two parts: first updating uncertainty models with a new extension of ensemble Kalman filter and second, feeding the updated uncertainty models to a neural network agent (trained using policy gradient reinforcement learning) that adapts the destination decisions of extracted material. The proposed framework is applied at a copper mining complex, which shows its applied aspects and an excellent performance to respond and integrate the new incoming information efficiently in an operational mining environment for adapting the materials flow. The proposed framework better meets the capacity and blending requirements of the different processing mills of the copper mining complex compared to the mine's cut-off grade destination policies. The proposed framework generates an additional 11%, 27%, 29%, and 29% of copper, gold, silver, and molybdenum products, respectively, and an additional 15% of cash flows, as compared to the mine's cut-off grade destination policies for the initial simulation. The extended ensemble Kalman filter updates multivariate local features of the mineral deposits with new blasthole information. The neural network destination policies are better at responding to the new information and adapt the destination decisions over the updated simulations more intelligently to meet the targets better. The updated destination decisions from neural network destination policies generate an additional 19%, 53%, 71%, and 76% of copper, gold, silver, and molybdenum products, respectively, as well as an additional 22% of cash flows. The mine's cut-off grade destination policies only generate an additional 8%, 38%, 56%, and 61% of copper, gold, silver, and molybdenum products, respectively, and an additional 11% of cash flows,

over the updated simulations. The proposed framework only adapts the destination decisions of the mining blocks, thus limiting the full potential and use of new information. In the future, a framework that can adapt all the relevant decisions of the short-term production plan will be developed.

**Acknowledgements** The work in this paper was funded by the National Sciences and Engineering Research Council (NSERC) of Canada CRD Grant 500414-16 and NSERC Discovery Grant 239019, the industry consortium members of McGill University's COSMO Stochastic Mine Planning Laboratory (AngloGold Ashanti, Barrick Gold, BHP, De Beers, IAMGOLD, Kinross Gold, Newmont Corporation, and Vale); and the Canada Research Chairs Program.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix: Notations

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

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

Parameters	Definition
$S$	Set of stochastic orebody simulations, $s \in S$
$\mathbb{S}$	Set of joint uncertainty scenarios that include orebody and equipment simulations, $s \in \mathbb{S}$
$\mathbb{T}$	Production planning horizon, $t \in \mathbb{T}$
$\mathcal{M}$	Set of mines in a mining complex, $m \in \mathcal{M}$
$\mathcal{B}_m$	Set of mining blocks in a mine $m$ , $b \in \mathcal{B}_m$
$\mathcal{N}$	Set of internal nodes in a mine
$x_i$	Location of internal nodes in a mining block $b$ , $i \in [1, \mathcal{N}]$
$\mathcal{C}$	Set of crushers in a mining complex
$\mathcal{P}$	Set of processing mills in a mining complex
$\mathcal{L}_O$	Set of oxide leach pads in a mining complex
$\mathcal{L}_S$	Set of sulphide leach pads in a mining complex
$\mathcal{W}$	Set of waste dumps in a mining complex
$t'$	Time step when new information is collected
$V$	Set of internal nodes in a mining block $b$
$E$	Set of elements in a mineral deposit, $e \in E$

**Table 2** continued

Parameters	Definition
$d_e^{t',s}(x_i)$	Initial data value of the internal node at the location $x_i$ for element $e$ at time $t'$ , and scenario $s$
$d_e^{t',s}(b)$	Initial data value of mining block $b$ for element $e$ at time $t'$ and scenario $s$
$\mathbb{Z}_e^{t',s}(x_i)$	Initial MAF value at location $x_i$ for element $e$ at time $t'$ and scenario $s$
$\mathbb{Z}_e^{t',s}(x)$	Vector of $\mathbb{Z}_e^{t',s}(x_i)$ for element $e$ at time $t'$ and scenario $s$
$A_{t'}$	Matrix of the contribution of internal nodes towards new information at the time $t'$
$A_{t'}^T$	Transpose of matrix $A_{t'}$
$\epsilon_e^{t'}$	Error in the new information for element $e$ at time $t'$
$P_e^{t',s}$	Model-based prediction for element $e$ at time $t'$ and scenario $s$
$K_e^{t'}$	Kalman gain for element $e$ at time $t'$
$C_{o_e o_e}^{t'}$	Measurement error covariance matrix for element $e$ at time $t'$
$C_{u_e u_e}^{t'}$	Model error covariance matrix for element $e$ at time $t'$
$u_e^{t',s}(x_i)$	Initial Gaussian values for element $e$ at time $t'$ generated by transforming MAF values at the location $x_i$
$U_e^{t',s}(x)$	Vector of $u_e^{t',s}(x_i)$ for elements $e$ at time $t'$
$m_e^{t'}$	MAF value of new information transformed to Gaussian values for element $e$ at time $t'$
$U_{a,i,t}$	Upper production limit for property $a$ at $i$ in period $t$
$L_{a,i,t}$	Lower production limit for property $a$ at $i$ in period $t$
$C_{a,i}$	Cost of processing material property $a$ at $i$
$c_{a,i}^+$	Cost of deviation from the upper target $U_{a,i,t}$ for material property $a$ at $i$
$c_{a,i}^-$	Cost of deviation from the lower target $L_{a,i,t}$ for material property $a$ at $i$
$nI$	Number of input neurons
$nH$	Number of hidden neurons
$nO$	Number of output neurons
$h_j$	Hidden neuron $j$
$o_k$	Output neuron $k$
$m_{b,s}$	Mass of block $b$ under scenario $s$
$g_{a,b,s}$	Grade of material property $a$ in block $b$ under scenario $s$
$\epsilon_e^{t'}$	Noise in the new information for element $e$ at time $t'$
$nIter$	Number of training iterations
$g_i$	Gradient at iteration $i$ , $i \in [1, nIter]$
$\eta$	Decay rate
$\partial$	Smoothing term
$\gamma$	Learning rate
$\Phi_M^e$	MAF transformation function for element $e$
$\Phi_G^e$	Gaussian transformation function for element $e$
$\Phi_M^{e^{-1}}$	MAF inverse transformation function for element $e$

**Table 2** continued

Parameters	Definition
$\Phi_G^{e^{-1}}$	Gaussian inverse transformation function for element $e$
$C_{u_e u_e}^{t'}$	Model error covariance matrix for element $e$ at time $t'$
$C_{o_e o_e}^{t'}$	Measurement error covariance matrix for element $e$ at time $t'$
$l_e^{t'}$	MAF value of new information collected for element $e$ at time $t'$
$\mathbb{P}_T$	Property tonnage that flows in the mining complex
$\mathbb{P}_M$	Set of metal properties that flow in the mining complex
$T_{m,s}^l$	Loading time with the shovel at mine $m$ under joint uncertainty scenario $s$

**Table 3** Variables used in the proposed framework

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

## References

- Aissani, N., Bekrar, A., Trentesaux, D., & Beldjilali, B. (2012). Dynamic scheduling for multi-site companies: A decisional approach based on reinforcement multi-agent learning. *Journal of Intelligent Manufacturing*, 23(6), 2513–2529. <https://doi.org/10.1007/s10845-011-0580-y>.
- Asad, M. W. A., Qureshi, M. A., & Jang, H. (2016). A review of cut-off grade policy models for open pit mining operations. *Resources Policy*, 49, 142–152. <https://doi.org/10.1016/j.resourpol.2016.05.005>.
- Barde, S. R. A., Yacout, S., & Shin, H. (2019). Optimal preventive maintenance policy based on reinforcement learning of a fleet of military trucks. *Journal of Intelligent Manufacturing*, 30(1), 147–161. <https://doi.org/10.1007/s10845-016-1237-7>.
- Benndorf, J. (2015). Making use of online production data: Sequential updating of mineral resource models. *Mathematical Geosciences*, 47(5), 547–563. <https://doi.org/10.1007/s11004-014-9561-y>.
- Benndorf, J., & 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. <https://doi.org/10.1080/14749009.2015.1107342>.
- Blom, M., Pearce, A. R., & Stuckey, P. J. (2019). Short-term planning for open pit mines: A review. *International Journal of Mining, Reclamation and Environment*, 33(5), 318–339. <https://doi.org/10.1080/17480930.2018.1448248>.
- Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of the COMPSTAT'2010* (pp. 177–186). [https://doi.org/10.1007/978-3-7908-2604-3\\_16](https://doi.org/10.1007/978-3-7908-2604-3_16).
- Brewer, A., Nancy, S., & Thomas, L. (1999). Intelligent tracking in manufacturing. *Journal of Intelligent Manufacturing*, 10(3), 245–250. <https://doi.org/10.1023/A:1008995707211>.
- Chen, Y., & Oliver, D. S. (2012). Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Mathematical Geosciences*, 44(1), 1–26. <https://doi.org/10.1007/s11004-011-9376-z>.
- Dalm, M., Buxton, M. W. N., & van Ruitenbeek, F. J. A. (2018). Ore–waste discrimination in epithermal deposits using near-infrared to short-wavelength infrared (NIR–SWIR) hyperspectral imagery. *Mathematical Geosciences*, 51(7), 1–27. <https://doi.org/10.1007/s11004-018-9758-6>.
- Dalm, M., Buxton, M. W. N., Van Ruitenbeek, F. J. A., & Voncken, J. H. L. (2014). Application of near-infrared spectroscopy to sensor based sorting of a porphyry copper ore. *Minerals Engineering*, 58, 7–16. <https://doi.org/10.1016/j.mineng.2013.12.016>.
- Desbarats, A. J., & Dimitrakopoulos, R. (2000). Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors. *Mathematical Geology*, 32(8), 919–942. <https://doi.org/10.1023/A:1007570402430>.
- Dimitrakopoulos, R., Farrelly, C. T., & Godoy, M. (2002). Moving forward from traditional optimisation: Grade uncertainty and risk effects in open pit design. *Mining Technology*, 111(1), 82–88. <https://doi.org/10.1179/mnt.2002.111.1.82>.
- Dimitrakopoulos, R., & Godoy, M. (2014). Grade control based on economic ore/waste classification functions and stochastic simulations: Examples, comparisons and applications. *Mining Technology*, 123(2), 90–106. <https://doi.org/10.1179/1743286314Y.0000000062>.
- Dimitrakopoulos, R., & 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. <https://doi.org/10.1023/B:MATG.0000037737.11615.df>.
- Dovera, L., & Della Rossa, E. (2011). Multimodal ensemble Kalman filtering using Gaussian mixture models. *Computational Geosciences*, 15(2), 307–323. <https://doi.org/10.1007/s10596-010-9205-3>.
- Evensen, G., Carlo, M., & 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., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the 13th international conference on artificial intelligence and statistics* (pp. 249–256). <http://proceedings.mlr.press/v9/glorot10a.html>
- Goetz, A. F. H., Curtiss, B., & Shiley, D. A. (2009). Rapid gangue mineral concentration measurement over conveyors by NIR reflectance spectroscopy. *Minerals Engineering*, 22(5), 490–499. <https://doi.org/10.1016/j.mineng.2008.12.013>.
- Goodfellow, R., & Dimitrakopoulos, R. (2016). Global optimization of open pit mining complexes with uncertainty. *Applied Soft Computing*, 40, 292–304. <https://doi.org/10.1016/j.asoc.2015.11.038>.
- Goodfellow, R., & Dimitrakopoulos, R. (2017). Simultaneous stochastic optimization of mining complexes and mineral value chains. *Mathematical Geosciences*, 49(3), 341–360. <https://doi.org/10.1007/s11004-017-9680-3>.
- Hinton, G. E., Srivastava, N., & Swersky, K. (2012). *Neural network for machine learning—Lecture 6a: Overview of mini-batch gradient descent*. Retrieved January 1, 2016, from [http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture\\_slides Lec6.pdf](http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides Lec6.pdf).
- Hou, J., Zhou, K., Zhang, X. S., Kang, X. D., & Xie, H. (2015). A review of closed-loop reservoir management. *Petroleum Science*, 12(1), 114–128. <https://doi.org/10.1007/s12182-014-0005-6>.
- Iyakwari, S., Glass, H. J., Rollinson, G. K., & Kowalczyk, P. B. (2016). Application of near infrared sensors to preconcentration of hydrothermally-formed copper ore. *Minerals Engineering*, 85, 148–167. <https://doi.org/10.1016/j.mineng.2015.10.020>.
- Jewbali, A., & Dimitrakopoulos, R. (2011). Implementation of conditional simulation by successive residuals. *Computers & Geosciences*, 37(2), 129–142. <https://doi.org/10.1016/j.cageo.2010.04.008>.
- Kargupta, H., Srakar, K., & Gilligan, M. (2010). MineFleet<sup>®</sup>: An overview of a widely adopted distributed vehicle performance data mining system. In *Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 37–46). <https://doi.org/10.1145/1835804.1835812>.
- Koellner, W. G., Brown, G. M., Rodríguez, J., Pontt, J., Cortés, P., & Miranda, H. (2004). Recent advances in mining haul trucks. *IEEE Transactions on Industrial Electronics*, 51(2), 321–329. <https://doi.org/10.1109/TIE.2004.825263>.
- Kumar, D., & Srinivasan, S. (2019). Ensemble-based assimilation of nonlinearly related dynamic data in reservoir models exhibiting non-Gaussian characteristics. *Mathematical Geosciences*, 51(1), 75–107. <https://doi.org/10.1007/s11004-018-9762-x>.
- Lamghari, A. (2017). Mine planning and oil field development: A survey and research potentials. *Mathematical Geosciences*, 49(3), 395–437. <https://doi.org/10.1007/s11004-017-9676-z>.
- Lane, K. F. (1984). Cutoff grades for two minerals. In *Proceedings of the 18th international symposium on application of computers and operations research in mineral the industries* (pp. 485–492).
- Lane, K. F. (1988). *The economic definition of ore: Cut-off grades in theory and practice*. London: Mining Journal Books Limited.
- Mai, N. L., Topal, E., Erten, O., & Sommerville, B. (2019). A new risk-based optimisation method for the iron ore production scheduling using stochastic integer programming. *Resources Policy*, 62, 571–579. <https://doi.org/10.1016/j.resourpol.2018.11.004>.
- Matamoros, M. E. V., & 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. <https://doi.org/10.1016/j.ejor.2016.05.050>.
- Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., et al. (2013). *Playing Atari with deep reinforcement learning*. arXiv preprint, 1312.5602. <http://arxiv.org/abs/1312.5602>.
- Montiel, L., & 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. <https://doi.org/10.1016/j.ejor.2015.05.002>.
- Montiel, L., & Dimitrakopoulos, R. (2017). A heuristic approach for the stochastic optimization of mine production schedules. *Journal of Heuristics*, 23(5), 397–415. <https://doi.org/10.1007/s10732-017-9349-6>.
- Montiel, L., & Dimitrakopoulos, R. (2018). Simultaneous stochastic: Optimization of production scheduling at Twin Creeks mining complex, Nevada. *Mining Engineering*, 70(12), 48–56.
- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning* (pp. 807–814).
- Nguyen, D., & Bui, X. (2015). A real-time regulation model in multi-agent decision support system for open pit mining. In *Proceedings of the 12th international symposium continuous surface mining-Aachen* (pp. 255–262). <https://doi.org/10.1007/978-3-319-12301-1>.
- Paduraru, C., & Dimitrakopoulos, R. (2018). Adaptive policies for short-term material flow optimization in a mining complex. *Mining Technology*, 127(1), 56–63. <https://doi.org/10.1080/14749009.2017.1341142>.
- Paduraru, C., & Dimitrakopoulos, R. (2019). Responding to new information in a mining complex: Fast mechanisms using machine learning. *Mining Technology*, 128(3), 129–142. <https://doi.org/10.1080/25726668.2019.1577596>.
- Panzeri, M., Della Rossa, E. L., Dovera, L., Riva, M., & Guadagnini, A. (2016). Integration of Markov mesh models and data assimilation techniques in complex reservoirs. *Computational Geosciences*, 20(3), 637–653. <https://doi.org/10.1007/s10596-015-9540-5>.
- Quigley, M., & Dimitrakopoulos, R. (2019). Incorporating geological and equipment performance uncertainty while optimizing short-term mine production schedules. *International Journal of Mining, Reclamation and Environment*. <https://doi.org/10.1080/17480930.2019.1658923>.
- Rendu, J.-M. (2014). *An introduction to cut-off grade estimation*. Englewood, CO: Society for Mining, Metallurgy & Exploration.
- Rosa, L., David, Valery, W., Wortley, M., Ozkocak, T., & Pike, M. (2007). The use of radio frequency ID tags to track ore in mining operations. In *Proceedings of the 33rd application of computers and operations research in the mineral industries* (pp. 601–606).
- Rossi, M. E., & Deutsch, C. V. (2013). *Mineral resource estimation*. New York: Springer. [https://doi.org/10.1007/978-1-4020-5717-5\\_1](https://doi.org/10.1007/978-1-4020-5717-5_1).
- Ruder, S. (2016). *An overview of gradient descent optimization algorithms*. arXiv preprint arXiv:1609.04747. <https://doi.org/10.1111/j.0006-341X.1999.00591.x>.
- Sarma, P., Durllofsky, L. J., Aziz, K., & Chen, W. H. (2006). Efficient real-time reservoir management using adjoint-based optimal control and model updating. *Computational Geosciences*, 10(1), 3–36. <https://doi.org/10.1007/s10596-005-9009-z>.
- Shirangi, M. G. (2017). *Advanced techniques for closed-loop reservoir optimization under uncertainty* (Doctoral dissertation). Stanford: Stanford University.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489. <https://doi.org/10.1038/nature16961>.
- Sutton, R. S., McAllester, D., Singh, S., & Mansour, Y. (2000). Policy gradient methods for reinforcement learning with function approximation. In *Proceedings of the advances in neural information processing systems* (pp. 1057–1063). <http://papers.nips.cc/paper/1713-policy-gradient-methods-for-reinforcement-learning-with-function-approximation.pdf>.
- Vargas-Guzmán, J. A., & Dimitrakopoulos, R. (2002). Conditional simulation of random fields by successive residuals. *Mathematical Geology*, 34(5), 597–611. <https://doi.org/10.1023/A:1016099029432>.
- Verly, G. (2005). Grade control classification of ore and waste: A critical review of estimation and simulation based procedures. *Mathematical Geology*, 37(5), 451–475. <https://doi.org/10.1007/s11004-005-6660-9>.
- Vo, H. X., & Durllofsky, 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. <https://doi.org/10.1007/s11004-014-9541-2>.
- Wambeke, T., & Benndorf, J. (2018). A study of the influence of measurement volume, blending ratios and sensor precision on real-time reconciliation of grade control models. *Mathematical Geosciences*, 50(7), 801–826. <https://doi.org/10.1007/s11004-018-9740-3>.
- Xu, T., & Hernández, J. G. (2019). Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble Kalman filter. *Advances in Water Resources*, 112, 106–123.
- Xue, L., & Zhang, D. (2014). A multimodel data assimilation framework via the ensemble Kalman filter. *Water Resources Research*, 50(5), 4197–4219. <https://doi.org/10.1002/2013WR014525>.
- Yüksel, C., Minnecker, C., Shishvan, M. S., Benndorf, J., & Buxton, M. (2018). Value of information introduced by a resource model updating framework. *Mathematical Geosciences*, 51(7), 1–19. <https://doi.org/10.1007/s11004-018-9770-x>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.