Article

Discrete Event Simulation for Machine-Learning Enabled Mine Production Control with Application to Gold Processing

Felipe Peña-Graf ¹ , Javier Órdenes ², Ryan Wilson ² and Alessandro Navarra ^{2,*}

¹ School of Engineering, Universidad Católica del Norte, Larrondo 1281, Coquimbo 1781421, Chile; felipe.pena@ucn.cl

² Department of Mining and Materials Engineering, Faculty of Engineering, McGill University, 3610 University Street, Montreal, QC H3A 0C5, Canada; javier.ordenes@mail.mcgill.ca (J.Ó.); ryan.wilson@mail.mcgill.ca (R.W.)

* Correspondence: alessandro.navarra@mcgill.ca

Abstract: Interdisciplinary barriers separating data scientists and geometallurgists have complicated systematic attempts to incorporate machine-learning into mine production management; however, experiences in excavating a vein-hosted gold deposit within the Alhué region of Chile have led to methodological advances, which is the subject of the current paper. These deposits are subject to several challenges, from increasing orebody complexity and decreasing gold grades to the significant geological uncertainty that is intrinsic to these systems. These challenges then translate to mineral processing, which is already dealing with increased environmental and technological constraints. Geological uncertainty causes stockout risks that can be mitigated by the approach that is developed within this paper, which features alternate operational modes and related control strategies. A digital twin framework based on discrete event simulation (DES) and a customized machine-learning (ML) model is proposed to incorporate geological variation into decision-making processes, including the setting of trigger point that induces mode changes. Sample calculations that were based on a simulated processing plant that was subject to mineralogical feed changes demonstrated that the framework is a valuable tool to evaluate and mitigate the potential risks to gold mineral processing performance.

Keywords: machine-learning; discrete event simulation; digital twin; random forest; gold processing; operational modes

Citation: Peña-Graf, F.; Órdenes, J.; Wilson, R.; Navarra, A. Discrete Event Simulation for Machine-Learning Enabled Mine Production Control with Application to Gold Processing. *Metals* **2022**, *12*, 225. <https://doi.org/10.3390/met12020225>

Academic Editor: Chris Aldrich

Received: 21 December 2021

Accepted: 21 January 2022

Published: 25 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Gold has been one of the most important and sought-after precious metals throughout human history. Its uses have ranged from jewelry, ornaments, and luxury items, to functions as money and capital reserves, and, more recently, as a main technological component benefitting modern-day living standards. Nevertheless, its extraction and beneficiation processes are becoming increasingly complex and are limited by several challenges.

Mining systems are designed and configured to obtain maximum profits that are bounded by technological limitations, environmental regulations, and the tactical restrictions that align operational short-term and strategic long-term objectives. This is particularly relevant in the gold mining industry which is currently faced with increasing complexities that are related to deposit locations and morphologies, decreasing gold grades, and difficult metallurgical profiles such as lower gold–sulfur ratios and higher proportions of impurities (i.e., penalty elements). These factors are especially impactful to deposits that operate with tight margins [1]. One of the main challenges confronting the gold mining industry is the geological uncertainty that is associated with increased variability and complexity of mineral distributions in newer discoveries. It is, therefore, critical that the strategic decisions to develop a mineral deposit take this geological uncertainty into account [2].

The often complex, erratic, and localized nature of gold is a common feature of many vein-style gold deposits [3] and can generally be linked to the underlying mineralogical distributions and downstream geometallurgical variables which are also heterogeneous. The uncertainty regarding ore grade, spatial distribution of the different host-rock lithologies, and the variability of metallurgical parameters are important aspects to consider during the mine planning process. Indeed, the incorporation of geological uncertainty into the block model (using computer aided design) is a main factor to determine how the deposit will ultimately be exploited. Moreover, the prediction of mining losses (as in room and pillar, and strip methods), depends on the incorporation of geological uncertainty within the control strategies and affects forecast production volumes and required adjustments of downstream processes. These challenges depend considerably on the genetic model of the deposit and the related spatial variability [4,5]. The geological uncertainty stems from our limited knowledge of the deposit due to finite sampling of the orebody and surrounding areas [3]. Thus, our knowledge of the correct geological model, deposit genesis, and ore characterization are also limited, which affects the ability to properly evaluate resource and reserve bases [5,6].

Along with the geological uncertainty, there could be potential significant mineralogical changes that could have important negative effects on the overall mine operation. This is due to the fact that a large portion of the sampling and testing prior to development takes place in the early stages of project evaluation with ore that is reachable through common methods such as surface drilling. Significant changes in ore characteristics are commonly observed both laterally and to depth; in other words, ore that is exploited in the early production stages (e.g., bulk samples) does not necessarily fully represent the characteristics of the entire deposit. Therefore, plants may face two main alternatives to respond to drastic mineralogical changes and related metallurgical responses: (1) make significant investment in assets to add new processing pathways, or (2) develop mineral blending strategies to achieve improved stability in ore feeds over an extended mine life period. Blending is useful to plan block extraction over a number of years; each block must be assigned to a destination process stream while also satisfying various constraints, such as the availability of operational resources (e.g., transportation and equipment) and maximum or minimum concentrations of target minerals or impurities [4]. Blending strategies that are paired with alternate modes of operation in the planning and development of mineral processing plants represent a viable alternative to maximize productivity.

The issues of aging mines that seek to maintain competitiveness are longstanding [6]. In recent years, however, there has been growing interest in incorporating machine-learning into operational decision-making, leading to concise treatments such as that of Burkov [7]. Such treatments are intended for broad appeal to industrial practitioners, as well as data scientists and researchers. However, these general treatments of machine-learning enabled control do not address the particularities of mining development, including geological uncertainty. Indeed, the underdevelopment of methodologies that incorporate machine-learning into mining control decisions can lead to confusion, which is effectively another form of uncertainty, i.e., “what can we do with these novel techniques at our mine site?”, “How might they be useful to us in practice at our mine?”, etc. which is confounded with the aforementioned geological uncertainty.

The advantage of a long-running mine is the practical experience that the local staff have accumulated over years and the deep knowledge that they have about their process and orebody. Following observations at their mine site, geometallurgists may form hypotheses of the type; the coordination of *these* aspects of the geology with *those* aspects of the metallurgical process would enhance the competitiveness of the mine. Prior to recent developments, the testing of these hypotheses might have been limited the fitting of linear and multilinear regressions, which can actually be considered as a form of machine-learning (See Section 2.1). Moreover, the comprehension of the physicochemical phenomena may benefit from a combination of theoretical and experimental analysis. However, a more complete testing of the geometallurgically-driven hypothesis would consider other machine-learning

techniques that can address the nonlinear and nested complexity that relates geological attributes to metallurgical extraction [7]. Furthermore, it remains unclear how a proposed machine-learning scheme will function when confronted with the geological variation of a given mine site unless it can be brought into an integrated simulation platform that mimics the mine site.

The objective of the current study is to reconcile the interdisciplinary divide between geometallurgists and data scientists by demonstrating a simulation methodology to formulate and evaluate the machine-learning-assisted characterization of geological uncertainty to be used in the development of integrated control strategies. The methodology builds on knowledge of machine-learning techniques, discrete event simulation, and digital twin development, as described in the following section, which each contribute to the subsequent sample computations (Figure 1). These computations were developed in collaboration with the Florida Mine (located in the Alhué district of Chile), demonstrating a random-forest machine-learning model that was applied to industrial data, and thus demonstrating the potential impact of the operational policies that are proposed. Following the computations there is a concluding discussion, highlighting the main aspects of the experience in developing the methodology and computations, as well as potential future work.

Figure 1. Concepts that were integrated into the contextualized sample calculations.

2. Methodology Development

2.1. Machine-Learning

Machine-learning is a subfield of computer and data science that is concerned with designing and building algorithms that are capable of ‘learning’ (i.e., improving its parameters) from a collection of examples of some observable phenomenon [7]. Figure 2 shows a commonly-applied classification scheme for different machine-learning models that, based on the data that are used to train the model, could be categorized as supervised, unsupervised, and reinforcement learning. Table 1 expands upon the structure of Figure 2 to include specific machine-learning models, some of which can be adapted for more than one of the categories of Figure 2. Moreover, the general models that are listed in Table 1 can be customized or adapted into particular models that are not explicitly mentioned in the table. Each of the models and approaches that are described in Table 1 have extensive literature, which can be overwhelming to practitioners within the mining industry. These techniques are recognized as having contributed to optimal resource allocation and process optimization in numerous industries, but it is usually unclear how these techniques may be combined to enhance specific mining operations. The current paper presents an approach to experiment with adaptations of these algorithms within a virtual representation (simulation) of a mine, prior to actual implementation. The divisions within Figure 2 and Table 1

are further explained below, prior to focussing on one particular machine-learning model, random forests.

Figure 2. The basic classification scheme for different types of machine-learning.

Table 1. A non-exhaustive list of machine-learning models and approaches that are categorized according to basic types of machine-learning.

Machine-Learning Model	Regression	Classification	Unsupervised Learning	Reinforcement Learning
Linear and multilinear regression	✓	-	-	-
Polynomial regression	✓	-	-	-
Logistic regression	✓	✓	-	-
Kernel regression	✓	-	-	-
Artificial neural network	✓	✓	-	-
Decision trees	✓	✓	-	-
Random forest	✓	✓	-	-
Support vector machine	-	✓	-	-
Naïve Bayes classifier	-	✓	-	-
K-means clustering	-	-	✓	-
Hierarchical clustering	-	-	✓	-
Mean shift clustering	-	-	✓	-
Density-based clustering	-	-	✓	-
Principal component analysis	✓	-	✓	-
Reinforcement learning based on Q-Learning	-	-	-	✓
Partially supervised reinforcement learning	✓	✓	-	✓

In supervised learning, the dataset is a collection of labeled examples, i.e., the dataset is organized to distinguish between attributes that are inputs and those that are outputs; in this case, the machine-learning algorithms are used to create a map that relates the inputs to the outputs. On the other hand, unsupervised learning uses a collection of unlabeled examples from which the algorithm aims to discover the underlying patterns in the data; this may involve partitioning the data into well-characterized clusters [7], noting that Table 1 lists four such clustering techniques. Thirdly, reinforcement learning describes a dynamic process, which may or may not benefit from an initially available set of data but is instead concerned with interacting with a system to progressively enhance its learnt understanding (parametrization) of the system.

As shown in Figure 2, supervised learning algorithms can be further subdivided into regression and classification models, which are regarded separately within Table 1. Whereas regression involves the generation of continuous target values, classification is focused on determining discrete outputs (classes). Both of these model types are trained based on a set of predictor input variables and corresponding output values. Certain approaches are routinely adapted for both purposes, although not all adaptations are included in Table 1, e.g., multilinear regression can be adapted for categorization, for example if the output variable is postprocessed to give a 0–1 indicator variable; however, this task is

more properly accomplished through logistic regression. The sample computations that are presented later in the paper (Section 3) focus on random forest classifiers, due to their relative simplicity and wide acceptance. Indeed, a random forest algorithm will be incorporated into a discrete event simulation (DES), to demonstrate the role of DES in the development of machine-learning-enabled digital twins.

However, the flexibility of the current framework would allow for the integration of other machine-learning techniques, depending on the specific application. Machine-learning and artificial intelligence are now at the forefront of modern automation technology, influencing almost every industry. As such, the utilization of such tools in the mining space are beginning to accelerate. Recent work covers mineral exploration, planning, operation, and mineral processing [8–10], although these works do not present the connection to digital twin control strategies. A number of recent reviews that are specific to mineral processing highlight the current state-of-the-art and research trends, and outline key areas of interest as a pathway forward in the mining industry. Ali and Frimpong [8] emphasize that future work should focus on generalizable models that can be applied to different settings or target commodities. McCoy and Auret [9] point out that machine-learning techniques are frequently applied as soft sensors for the prediction of difficult or infrequently measured variables but are often not integrated with quantitative frameworks that could be fed with those measurements. Finally, Cisternas et al. [10] stressed the importance of integrating detailed optimization models, e.g., at the molecular level, into higher level models to improve overall system performance. Our work is well-placed within the current literature by presenting a holistic and flexible framework for mine production control that lowers the entry barriers for industrial-scale design and implementation.

The remainder of this section focusses on random forests, in anticipation of their usage within Section 3. A random forest is an ensemble-supervised machine-learning model that is used for classification and regression tasks by combining multiple decision trees [11]. These decision trees depend on the values of a random vector that is sampled independently from the set of features and with the same distribution for every tree in a method that is known as bootstrap aggregating (a.k.a. bagging) [12]. Within each tree, splits are generated to maximize the decrease of impurity that is introduced by a split [13]; the so-called “impurity” refers to the likelihood of incorrectly classifying a randomly chosen element in the dataset if it were given a random label based on the class distribution.

The random forest algorithm is resistant to problems of overfitting, and, therefore, performs comparatively in the presence of noisy data [13]. Moreover, its simplicity and relatively easy implementation makes it a popular algorithm that has been widely used in several fields for classification and regression problems [14].

In a classification model, decisions are made by averaging the class assignment probabilities that are produced by every tree. Prediction of a new datapoint is produced by evaluating every decision tree in the forest, with each casting a class membership vote, as illustrated in Figure 3. The final membership class is decided by plurality vote.

Normally, the main parameters that are set to implement the algorithm are the number of trees and the number of features that are used in each. The number of trees (each with low bias and high variance) determines the size of the forest [11], and the number of features determines the number of variables to select and test for the best split in each new tree instance. The value that is used for the number of trees could be in the range of 100–500 [15]; too far beyond this number of trees typically results in diminishing returns [16]. The number of variables is generally equal to the square root of the number of inputs [17].

Due to the difficulty in interpreting the forest of trees that are generated, variable importance and selection methods are used to understand the interactions of the input variables on the output results. In particular, Gini importance, also known as impurity importance, is calculated for each variable as the sum of all decreases in the impurity measure at all nodes in the forest at which a split on the selected variable has occurred [13].

To evaluate model performance, the two most frequently used metrics are precision and recall. Precision is defined as the ratio of correct positive predictions to the overall number of positive predictions, meanwhile recall is defined as the ratio of correct positive predictions to the overall number of positive examples in the test set [7]. To complement the results, additional receiver operating characteristics (ROC) curves can be plotted to evaluate the classification model and visualize its performance by depicting the relative trade-offs between benefits and costs [18].

Figure 3. Generalized diagram for the functioning of a random forest classifier.

2.2. Discrete Event Simulation

Discrete event simulation (DES) is a computational paradigm to develop dynamic models that include the interaction of critical variables and processes of discrete event systems. These systems are characterized by a discrete state space in which its state evolution depends on asynchronous discrete events over time [19]. In this discrete space, objects (a.k.a., entities) are represented individually and can be tracked through the system. Specific attributes are assigned to each individual object and determine what happens to them throughout the simulation. Many of the earliest applications of DES that were related to queuing networks, such as in job shops and other manufacturing contexts, in which the attributes describe the objects (entities) and the processes that transform these objects. Due to its discrete nature, state changes occur at discrete points in time. In recent years, its use and capabilities have been enhanced with the appearance of fast and inexpensive computing capacity [20].

Computer-based DES is a powerful design and testing tool to support decision-making within industrial systems by simulating thousands of operational days to evaluate the impact, bottlenecks, risks, and other deficiencies of operational policies. DES frameworks can then be used to develop digital twins to design and test operational policies at all stages of the mining life cycle [21], especially within multiphase engineering projects and continuous improvement initiatives [22].

One of the key differences between mining and other industrial systems is that mining operations are subject to geological uncertainty. In other words, due to the natural variability of orebodies and host geological environments, there is uncertainty as to whether actual ore feed characteristics will vary from the expected composition. Each deposit may have a minimum grade (i.e., the ore cut-off grade), and other criteria that would determine whether sections of it can be profitably excavated and processed. As DES offers the flexibility to incorporate random distributions, the geological uncertainty (and its economic implications) can be modelled and mitigated by the implementation of operational buffers, such as flexible blending strategies, stockpiles, pre-treatment processes, and alternative operational modes.

In the mining industry, DES is used to conduct “what-if” analysis by supporting mining engineers and management in decision-making [23], particularly in the analysis of operational buffers. Indeed, DES frameworks have been applied to material management, equipment selection, transportation, and continuous mine system simulation. Other studies have explored DES applications for mine–mill modelling [24]; these include the use of DES to quantify the impact of ore-type spectral imagery, the use of solar power on semi-autogenous grinding (SAG) mills, and stochastically predicted ore production based on a truck haulage system [25]. More recently, work by Navarra et al. [26] opened a new research area regarding mineral processing and simulation through mass balance and mathematical programming by proposing the use of alternate operational modes. This approach combined with DES has been applied in several areas of the mining industry, such as concentrator and smelter dynamics [27–29], heap leach processes [30], chromite beneficiation modeling [22], reagent consumption in gold processing [31], oil sands processing [21], and tailing retreatment applications [24].

Currently, areas such as manufacturing [32] and healthcare [33] have been adopting new strategies to integrate ML models into DES frameworks to produce better results and more realistic representations. More precise representations of the process variability and the consequences of operational decisions can help develop better control strategies [31,33]. Greasley [34] reviews several different architectures that implement commercial off-the-shelf (COTS) DES and machine-learning models, which are summarized in Figure 4. In the offline category (Figure 4a), the interaction between DES and ML occurs via data files that contain the output of one of the models, which later serves as the input for the other. Conversely, the online category (Figure 4b) integrates both models through a data interface that does not require human intervention; communication can occur directly between software programs, or via servers and/or the internet.

Figure 4. Various architectures for combining COTS DES and ML software, classified as (a) offline, and (b) online. (adapted from Greasley [34]).

Based on these architectures and their interactions, ML strategies can serve to model input data for a DES framework, or outright give the instructions to build the DES model, providing a digital reality [35]. This digital reality that is generated through data-driven methods enables reconfigurations of the simulation model to reflect the actual state of the system as digital twins. Similarly, DES can serve to model input data and serve as an experimentation and analysis tool for subsequent ML modeling [35]. This study develops an offline model with an architecture that develops a DES through a ML-assisted ore classification.

2.3. Digital Twin Development

To model complex situations, such as geological uncertainty in mining systems, operational models must be able to represent and respond to the most impactful aspects of the operational variability. Recent work by Navarra et al. [28] developed a DES framework to provide a representation of feed stockpiles via mass balancing to assess operational risk caused by unexpected changes in ore feed attributes. The risk of stockout is then mitigated by the alternation of operational modes with different blending strategies.

This version of the model is fed with sampled data derived from theoretical distributions based on the expected behaviour and incorporating a controlled level of uncertainty (e.g., standard deviation) through random number generation. Though this Monte Carlo approach allows for a good estimation in the design stages of a processing plant, later stages could benefit from increased geometallurgical and operational data by incorporating other methods such as geostatistical modeling [24] or predictive ML models [21] to develop the input data.

In this context, DES models can be expanded and consolidated to develop so-called digital twins. A digital twin can be defined as an integrated simulation of a system that uses the best available data models, updates, and history to mirror the behaviour of the corresponding physical system [35]. However, there are a variety of different types of digital twins depending on the specific application and related classification scheme [36]. For instance, a digital twin can be classified based on its level of integration. Whereas offline twins (a.k.a., digital models or shadows) depend on manual data exchanges between the physical and digital systems, full online implementations rely on sensors and other control systems to maintain automatic bidirectional data flow [36].

The concept of the digital twin is often credited to Grieves' 2002 work that is related to product lifecycle management (then coined 'Mirrored Spaces Model') [37], with the now widely used term 'digital twin' used by NASA in 2010 [38]. However, the general approach was actually pioneered by NASA's aerospace program in the 1960s [39]. Regardless, research interest in digital twins and their use in industrial settings have accelerated over the last two decades. In fact, some estimates indicate the digital twin market could be worth as much as nearly US \$50 billion by 2026 [40]. Despite this, the use of digital twins in the mining industry remains fairly limited to date, although interest and work in this area are rapidly gaining interest.

While certain offline digital twin solutions have been offered by service providers for several years, real-time online digital twins in the mining space have only begun to emerge in the last three to five years, and mostly remain under development. Early examples include the creation of digital twins by Andritz Automation for the cyclone feed line and underflow/overflow processes at OceanaGold's Haile mine in South Carolina (2017) [41], and by Petra Data Science for monitoring the effects of blasting parameters (e.g., blasthole spacing) on mill throughput at PanAust Ltd.'s Ban Houayxai mine in Laos (2018) [42]. Nonetheless, research and development of digital twin applications within the mining industry remain very much ongoing, particularly for mineral processing and related mine-to-mill integration.

To this end, a combined framework implementing machine-learning and discrete event simulation could be a powerful tool to stabilize mineral processing plant performance in response to predicted geological uncertainty earlier in the value chain. The

architectures that were described by Greasley [34] (Figure 4) provide some insight into digital twin development. Traditionally, the inputs are generated by sampling data from rules and theoretical distributions that are derived from domain knowledge, as was already mentioned. As an extension, this study proposes a machine-learning model that is based on random forest classification to generate data that is subsequently consumed by the DES framework model of Navarra et al. [28]. A similar approach was used by Wilson et al. [21] to create a digital twin which integrated partial least squares regression modelling into DES to model oil sands processing.

The high-level structure of the proposed methodology for digital twin development is represented in Figure 5, wherein real sampled data are passed into a predictive classification model to generate the proper inputs into the DES framework. The final output of the DES framework, which considers the geological variability that is inherent to mineralized systems, could then assist with the decision-making process in the design stages of production planning. In this form, the digital twin is initially intended for offline use; however, as plant development progresses and operational data accumulates, the framework can eventually be incorporated into the control system of the mine, to provide real-time feedback. It is important this transition occurs gradually, beginning with one-way automatic data transfer from the physical system to the digital twin, to ensure the appropriate functioning of the virtual framework. At this step, data flowing from the physical system (e.g., assay, mineralogical, or sensor data) goes directly to the data collection process of the digital twin. With sufficient validation and proof-of-concept, the digital twin can then be brought fully online to support a two-way automatic data exchange; such a real-time feedback loop has the advantage of simultaneously performing forward predictions and backward reconciliations to optimize the system parameters. DES can thus be an important avenue to develop digital twins for a variety of applications within dynamic mining systems.

Figure 5. Proposed methodology for digital twin development, in which data flows that interconnect the real and virtual control systems are initially offline (i.e., through data file transfer). Upon further development, a portion of the data flows are brought online (i.e., through a data interface).

The current approach proposes data flows between the physical system and the digital twin, via their respective control systems (Figure 5) that are initially offline, i.e., utilizing data files (Figure 4a). This data informs the adjustments of the real and virtual control systems. In subsequent development, the transmission of data and instructions can be brought online through the development of data interfaces (Figure 4b). The following section provides sample calculations in the context of a vein-hosted gold deposit. The comprehension of the real mining system (Figure 5, left side) requires inflows of the most critical data, which is site-specific; in some cases, for instance, the ore dilution at the stage of transportation (loading and hauling) might be critical, in which case this data should be incorporated into the control system.

It should be noted, however, that the control strategies are constrained by the underlying technological infrastructure. This is especially true for underground autonomous mining equipment, which can intermittently lose communication with the overarching control system [43]; indeed, early adopters of automated hauling systems (AHS) such as Codelco, Rio Tinto, and others, reported that unreliable data transmission and the resulting operational issues were a main reason preventing their wider adoption at more of their sites [44]. Following these earlier efforts, there have been advances in cellular technology, including high-performance protocols such as 5G, which will have increasing importance within mine control strategies. The communication framework is effectively the technological infrastructure for digital twin development. In the context of a development project, an offline simulation (Figure 4a) may be extended or adapted to actively inform the online control system (Figure 4b), but may be limited by the communications infrastructure, including routers, load balancers, cloud computing resources [45], etc. that are becoming increasingly prevalent in industrial control systems. The implementation of an online digital twin may, therefore, be part of a larger project, which includes the simultaneous upgrading of the infrastructure and the control strategies; a similar discussion has been described in the context of smelter operations [27] and has a broader relevance throughout the minerals industry.

Figure 5, in relation to Figure 4, is illustrative of our general approach to digital twin development. It is a general guide to first develop an offline simulation that incorporates control mechanisms explicitly within the simulation; DES models that follow this approach can then be extended/adapted into the real-time control system of a functional mine, but this may require a simultaneous upgrade of the of the underlying infrastructure [27,45]. This is distinguished from other applications of DES that do not explicitly simulate the control mechanism and are, therefore, not directly related to digital twin development. Moreover, the approach of Figure 5 allows the testing and refinement of machine-learning-enabled control strategies, prior to their deployment in an actual mine. It should be noted, however, that Figure 5 is intentionally general and its application requires context, e.g., the following section adapts our general approach for gold production using real industrial data from the Alhué district to parameterize machine-learning-enabled control strategies.

3. Sample Calculations

To demonstrate the overall framework, sample computations consider a generic vein-hosted gold deposit that is loosely based on mineralogical changes that are observed at several type examples that are located in Chile. In general, vein-hosted gold deposits are closely related with fault systems, dilatational fractures, or other structures which act as fluid conduits for large volumes of hydrothermal fluids and structural traps to concentrate the precious metal [46]. Several vein-hosted gold deposits have different genetic models, such as orogenic, epithermal (high and low sulphidation), Carlin-type, and turbidite-hosted, among others. Vein-type gold deposits are characterized by the sizes and geometries of orebodies, mineralization styles, and the deformation history of vein-forming events. Most mineralized zones occur within or proximal to shear zones, particularly those in larger systems with intersecting structure sets [47], which can generate tabular-shaped ore shoots.

Such deposits can extend on the order of hundreds of meters along two dimensions with substantially less breadth along the third minor dimension [48].

An important factor to consider is the interaction of an orebody with other geological features, such as the water table, topography, emplacement depth, uplift/exhumation processes, and younger faults. Any such feature can result in post-depositional mineralogical changes which can lead to significant geological variability and related uncertainty for a given orebody. A good example of this type of interaction exists in the Alhué mining district, where the distribution of metallurgical domains is strongly influenced by (1) the localized topography, and (2) the permeability of the host rock as a function of the major fault systems that interact with the veins. Both factors affect the location of the paleo-water table, which conditions the nature of mineral–groundwater interactions, and thus, the in-situ oxidation-reduction state of the minerals [31].

Other prime examples of such interactions include some of the high sulphidation epithermal deposits from the historic El Indio gold belt, which is located high in the Andes of north-central Chile. There are various styles of gold mineralization in the district, including vein-type (El Indio and Sancarrón), fault and hydrothermal breccias (Tambo), and strata bound deposits (Pascua). The El Indio deposit is characterized by sulphide-rich (hypogene) vein mineralization that is dominated by either copper (mainly enargite) or gold (native, auricuprite, calaverite, krennerite) [49,50]; in contrast, the Tambo and Pascua deposits have been variably oxidized [51,52]. While all three deposits are hosted by relatively permeable ignimbrites [52], Tambo and Pascua are interpreted to have been emplaced at shallower levels which would have exposed them to supergene oxidation and enrichment via near-surface weathering processes.

Similarly, Jerónimo (a.k.a. El Hueso), a manto-shaped strata-bound gold-rich carbonate replacement deposit that is located in the Atacama region of northern Chile, contains both oxide and sulphide-mineralized zones [53]. The oxidized zone (mined out as of 2002) consisted of native gold and electrum that was accompanied by manganese and ferric oxides; the sulphidic zone is dominated by pyrite $\pm$ pyrrhotite with lesser sphalerite, galena, bournonite, and pyrargyrite, and is considered refractory in nature [53,54].

Significant mineralogical changes can also be observed within an orebody strictly based on spatio-temporal (and related geochemical) differences between the mineralized zones. For instance, the low sulphidation El Peñón Au-Ag vein deposit has recently shown drastic changes in Ag and base metal mineral assemblages. The ore minerals that were observed in the earliest mined veins consisted of electrum (mostly 40–60 wt% Au), acanthite, native gold, native silver, and silver halides with rare pyrite, chalcopyrite, and galena. However, newly discovered veins are mainly composed of Ag sulfosalts (e.g., freibergite, pyrargyrite, polybasite) and sulfides (e.g., acanthite and argentite), which has caused a decrease in silver recovery in the plant [55].

It is clear from these few examples that vein-type precious metal deposits are often subject to significant variability between the various mineralized zones. As a result, it is not uncommon that an Au-Ag operation will experience substantial mineralogical changes in ore feeds over an extended production life. These changes can result in important consequences for the metallurgical profile and related process streams in the mineral processing plant. In the Chilean context, where many deposits possess appreciable supergene profiles due to the (semi-)arid climatic and geomorphological conditions, a common result is the division between oxidic and sulphidic ore domains. This causes mineral process engineers to consider multiple different mineral processing flowsheets, especially in the early stages of economic assessment for a mining project. In this paper, we will analyze the effects of a transition from oxide-dominant ore (<15% primary sulphide) to sulphide-rich ore on a direct leaching flowsheet assembly (originally intended for oxide ore).

The transition from low to high sulphide ore assemblages can severely affect the original metallurgical process when there is no profitable option to add a flotation circuit. In this context, a suitable approach might be to implement ore blending practices in tandem with alternate modes of operation to maximize the plant throughput and stabilize the

main process parameters. This study highlights the implementation of a digital twin framework using DES as a suitable tool to achieve these goals and mitigate the operational risk that is related to significant mineralogical changes in ore feeds. The framework is also shown as critical to the evaluation and selection of process alternatives, and thus towards the improvement of future production planning. Moreover, by integrating a predictive machine-learning algorithm into DES, the digital twin is capable of reducing the uncertainty in metallurgical gold and silver recoveries, in addition to better controlling reagent consumption (e.g., sodium cyanide) [31].

3.1. Input Data

To correctly assess the geological variation and mineralogical changes in the case that is presented, input data corresponding to a series of samples that were anonymized from the Alhué mining district was used. The data were collected from bottle roll testing for cyanidation samples, a common method for the evaluation of low-grade gold ores [56]. Descriptive statistics for the 2678 entries in the dataset are summarized in Table 2. In the proposed methodology, illustrated in Figure 5, the collection of such data is the first task within the digital twin development box; these data are then processed in the parameterization of the ML model.

Table 2. Descriptive statistics of the bottle roll test parameters that were used as input data (n = 2678 samples).

Variable	Description	Mean	Standard Deviation	Min	Max
lau	Gold Grade (g/t)	6.253	6.173	0.200	107.57
lag	Silver Grade (g/t)	140.148	223.272	1.9	5009.4
lcu	Copper Grade (ppm)	675.7380	1307.071	0	26,030.9
recau	Gold Recovery (%)	91.5	8.4	2.3	99.9
recag	Silver Recovery (%)	68.2	23.9	0	99.8
reccu	Copper Recovery (%)	37.1	22.5	0.5	100.0
cnc	Cyanide Consumption (kg CN/t ore)	2.706	1.646	0.009	9

For demonstrative purposes, the current work features the simulated use of a RF; however, the approach is general and could consider other ML models (Table 1) as alternatives to the RF, as well as complex approaches that combine several ML models within the control strategy. However, in all cases, an initial data analysis can help determine if an approach should be discarded even prior to attempting the simulation.

In the current example, the input data includes grades and recoveries for three potential pay-metals: gold, silver, and copper. Ideally this would be sufficient to estimate the minimum profitability criteria, assuming processing costs and metal prices. Typical cut-off grades in Chilean gold mines would be 0.5 g/t to 5 g/t of gold, and this depends on many factors, including the tradeable gold price and whether the gold is refractory (i.e., resistant to conventional processing techniques). The case of Alhué tends to be in the range of 3 g/t, but it is not so simple since copper sulphides (often considered to be a pay-metal) have a negative effect on the process, causing spikes in cyanide consumption; indeed, certain copper sulfides are said to be “cyanicidal” [31], i.e., they react and “kill” the cyanide thus prevent the desired interaction with gold. The final entry in Table 2 is the cyanide consumption, which is an especially critical concern in gold processing.

3.2. RF Model

As mentioned earlier, the goal of this research is to develop a digital twin that successfully represents a processing plant experiencing important mineralogical changes in ore feed. To improve the quality of the input data DES framework, a classification algorithm that is based on random forest is implemented to assess the sample destination; that is, whether the sample corresponds to a subset of high or low sulphide content. All the

predictor variables (Table 2) that were used and model training was conducted on 80% of the dataset at a time using five-fold cross-validation via the Python library Scikit-Learn [57]. A grid search optimization that was based on the area under the curve of the receiver operating characteristic (AUC-ROC) method [18], was performed to obtain the best training parameters. The final model parameters for the random forest model included 500 trees and a feature number that was equal to the square root of the number of variables (i.e., $7^{0.5} \approx 2.645$). Model metrics that were computed following the testing phase are presented in Table 3.

Table 3. Resulting classification metrics of the model.

Ore Category	Precision	Recall	F1-Score
Low Sulphide	0.79	0.75	0.76
High Sulphide	0.83	0.86	0.85

The overall accuracy of the model based on the test set was around 82% with an AUC of 0.88; the corresponding ROC curve is presented in Figure 6.

Figure 6. AUC-ROC curve of the RF Model compared with a no-skill classifier.

Results from a feature importance analysis, a measure of the impact of the input variables in the model based on Gini impurity [11], are presented in Figure 7. The results confirm the importance of variables such as Cu grade and Ag recovery, as previously reported by Órdenes et al. [31].

Figure 7. Feature importance plot of each variable resulting from the RF model.

3.3. DES Digital Twin

The following DES implementation is based on Navarra's simulation framework for processing plants in which two different ore types (Ore 1 and Ore 2) are blended and managed in alternative operational modes (Mode A and Mode B). In this case Ore 1 corresponds to a high sulphide ore and Ore 2 to a low sulphide ore. Recall that the low sulphide ore is the more desirable of the two since the conceptual plant was originally configured for it but is now limited by mineralogical changes in the extracted ore. Following a data-driven approach, geological variation is generated in the ore classification step (see Figure 5) to be used as an input for the DES simulations and represents the main source of uncertainty. The geological variation is then represented as follows:

- Input data for the DES model was obtained by a bootstrapping process on the original dataset (see Table 2). The bootstrapping process enforced a sample proportion of 55% high sulphide ore and 45% low sulphide ore for the study demonstration purposes.
- The bootstrapped data was then classified by the ML model and grouped to generate parcels of mined rock (i.e., collections of simulated mining blocks that are excavated, [28]) based on destination (40 samples for each parcel).
- The ore type proportions of the parcels were defined by the corresponding proportions that were obtained from the binary classification of the RF.
- The parcel sizes were randomly generated following a normal distribution with a mean of 40 kt of ore and standard deviation of 3 kt.

Within the terminology of DES (that is historically linked to the manufacturing context), the mineral process has the so-called attribute of “operational mode”, and the incoming parcel has the attributes of “ore grade” and “total mass”, which are implemented as dynamic state variables. Moreover, the mineral process that is based on sample data has higher variability and standard deviation from the controlled environment of previous implementations of the DES framework [21,22,31], which used a similar logic (Figure 8). Due to the source the data that were used, the simulations from this study also exhibited higher variability than in the earlier work [21,22,31], and thus considerably more stock variation. To account for this increased uncertainty, 100 statistical replicates were generated through this process.

Figure 8. High level representation of the DES flow model.

Figure 8 illustrates the main discrete events and decision points in the current DES implementation. Beginning with the left side of the figure, a mass of mined material that is called a “parcel” is simulated according to the random forest algorithm; the data for this first parcel is a component of the initialization. The parcel is mined and processed according to the active operational mode (in the center-left of Figure 8), until a discrete event is triggered, at which point a series of if-statements and state adjustments are executed (which constitute the entire right side of Figure 8). The event can be either the depletion of an ore stockpile (top diamond), the crossing of a critical threshold that would signify the need to change operational modes (middle diamond), or the depletion of the current parcel (bottom diamond).

For example, if the plant is currently operating under Mode A and Ore 2 is depleted below the set threshold, the system will switch to a contingency mode for a day to recover stock, after which regular Mode A operation is resumed. Moreover, if the Ore 2 level is below a critical threshold at the end of a production campaign, the decision is made to switch operational modes (in this example, to Mode B) during the shutdown period. The bottom diamond of Figure 8 shows that when a parcel is emptied (i.e., mined out), the system then proceeds to the next ML-characterized parcel under the active operational mode. Within this scheme, the discrete events include the depletion of stockpiles, the alternation between production campaigns and shutdowns, and the depletion of parcels. In this case, the shutdown times are predetermined, but further development of the framework can also include untimed disruptions and/or unforeseen circumstances as additional sources of uncertainty. The logic of Figure 8 was implemented and the results were generated using the commercial-off-the-shelf DES software Rockwell Arena®, with Visual Basic for Applications (VBA) extensions considering throughput as the main performance indicator. The current implementation considers an entity that represents the state of a mineral processing plant, and that manages the state of the parcel as it traverses through a time-consuming processes that advances the simulated clock; more details are provided in [28].

Table 4 describes the characteristics of the operational modes that were used for the current set of calculations. Mode A is a comparatively productive operational state that exhibits roughly the same throughput of the processing plant before encountering the mineralogical change, while Mode B is considered more stable and allows for stockpile replenishment and is particularly designed for this purpose based on bench-scale experiments. Both operational modes consider blends of Ore 1 and Ore 2 in different proportions. Considering the plant is better adapted to low sulphide ore, Mode A utilizes a larger proportion of Ore 2 compared to Mode B.

Table 4. Description of operational modes in relation to deposit forecast.

Operational Mode		Throughput (t/day)	Ore 1 in Feed (%)	Ore 2 in Feed (%)
Mode A	Regular	4300	45	55
	Contingency	2795	100	0
Mode B	Regular	3655	70	30
	Contingency	1827.5	0	100
Deposit	-	-	55	45

Based on this information, a deterministic analysis of operational modes can be implemented using mass balance equations that were developed by Navarra et al. [28] to define a steady-state optimum of the model. The results from these equations indicate that Mode A should be applied 1.28 times as often as Mode B, with an average output of 4016.48 t/day.

The framework is designed such that the processing plant acts as a bottleneck with mining rates exceeding the plant capacity. As such, campaign cycles are configured in a regular schedule of 29 days of production followed by a one-day shutdown; decisions on whether to change the operational modes are made during the shutdown phase. If the level of Ore 2 is low enough to present a potential stockout risk during the next campaign, a mode change is triggered from Mode A to Mode B. On the other hand, if Ore 2 levels

are above a defined threshold following a campaign in Mode B, the operational mode reverts to Mode A. In either case, should a stock shortage occur during an operating campaign, contingency modes are applied with a one-day timeframe. A contingency mode implies that the plant only utilizes the available ore type (refer to Table 3). Figure 9 depicts the day-by-day evolution of a simulation scenario and the fluctuation of the stockpiles of Ore 1 and Ore 2 under the respective operational mode. Note that around day 1340 and day 1360 the stock of Ore 2 entered a stockout period that triggered a contingency segment until the shutdown where the replenishing Mode B was scheduled for the next production campaign.

Figure 9. Stockpile levels of Ore 1 and Ore 2 over time during a simulation and the respective active operational mode, in which “A-Ac” denotes the alternation between the Regular and Contingency configurations of Mode A.

A simplified (deterministic) analysis of the optimal throughput of the simulated processing plant considers two operational policy parameters to characterize decision-making following the approach of Navarra et al. [28]:

$$\text{TOSL} = \text{Target Ore Stockpile Level} \quad (1)$$

$$\text{COSL} = \text{Critical Ore 2 Stockpile Level} \quad (2)$$

The target ore stockpile level is computed as the sum of the available ore types and the critical Ore 2 stockpile level corresponds to availability of the low sulphide ore.

To determine the potential impact of the new operational policies in response to the mineral transition, 28 scenarios have been tested with different values in both policy parameters, which are described in Tables 5 and 6. All of the TOSL and COSL values are expressed in kilotons denoted “kt”; in practice at Alhué, this is a useful unit since errors in daily estimates are in the range of 1 or 2 kt (these are verified through so-called metallurgical accounting that is based on statistical data reconciliation [58]). In the case of COSL, a range between 1 kt and 12 kt is considered while TOSL consider four different values for the stockpile of 10 kt to 40 kt. Each tested scenario corresponds to a simulation of approximately 1500 days of operation using the 100 replicas that were generated in the previous step. Table 5 summarizes the average daily throughput in tons of every scenario. Table 6 summarizes the average proportion of time in Mode A (highly productive) vs Mode B (replenishing mode).

Table 5. Summary of the average simulation throughputs in tons/day.

Target Stockpile Level (kt)	Critical Ore Stockpile Level (kt)						
	1	2	4	6	8	10	12
10	3812.7	3789.9	3780.87	3765.58	3754.49	3053.58	0
20	3948.52	3954.1	3956.39	3952.83	3942.59	3928.34	3907.7
30	3957.59	3965.38	3969.15	3972.26	3975.37	3976.71	3977.11
40	3957.33	3964.79	3969.23	3972.87	3976.1	3978.46	3980.6

Table 6. Summary proportion of the time spent in Mode A vs. Mode B for each scenario.

Target Stockpile Level (kt)	Critical Ore Stockpile Level (kt)						
	1	2	4	6	8	10	12
10	1.407	1.138	1.058	0.998	0.955	0	-
20	1.728	1.479	1.365	1.24	1.14	1.055	0.961
30	1.755	1.508	1.394	1.3	1.226	1.156	1.11
40	1.747	1.517	1.386	1.309	1.235	1.172	1.117

The simulation results exhibit throughputs nearing, but still below, the deterministic optimum. This is despite some scenarios spending proportions of time in Mode A vs. Mode B that were greater than the defined optimum. This behavior can be qualitatively explained by stockout periods during campaigns that force contingency modes which are unavoidably inefficient. To the extent of the sample computations, an optimal combination of TOSL and COSL should be found around 20 kt and 4–6 kt, respectively, and can be quantitatively pinpointed using simulation-based optimization tools such as OptQuest (see [28]). Nevertheless, economical, technical, and environmental variables could also be included to extend the scope of the framework as a risk assessment and decision-making tool based on a different set of key performance indicators (KPIs). It is important to note that lower control variable thresholds put a greater strain on the model and cause inefficiencies, particularly for systems that are dealing with significant geological variability and related uncertainty. On the other hand, excessively high threshold values (especially TOSL) may produce increased throughputs but will ultimately result in system instabilities due to spatial, economic, and environmental considerations. Future quantitative efforts can demonstrate the balancing of these potentially competing KPIs.

Nonetheless, the current results of the RF model, work sufficiently well to represent the uncertainty that is associated with mineralogical changes in a deposit with a more realistic distribution than a simplified Monte Carlo approach. Furthermore, the selected input data could be easily replaced by cheaper and quicker alternatives, such as shaker-table test data, that would allow for the predictive modelling of ore feed variability earlier in the mineral value chain. It is noteworthy, however, that such analytical replacements must only be undertaken once sufficient geometallurgical characterization data are available to build a truly robust predictive model. In this, the feature importance tests the association of Cu with a high sulphide ore, which also corresponds to poorer plant performance; this is consistent with previous findings that were reported by Ordenes et al. [31].

The overall results of this integrated digital twin framework demonstrate its potential to evaluate system-wide plant performance and the related risk factors under different operational modes and policies. By utilizing routinely collected data types, this presents a low-cost solution to support decision-making processes in response to geological variation within the mine–mill profile.

4. Discussion

Geological uncertainties are inherent to all ore deposits and, as such, all downstream mineral processing tasks are subject to variability and unexpected changes in ore feed attributes. This puts an important burden on integrated mine–mill profile management

as a deposit is developed, particularly for mature operations which may face significant geometallurgical changes. The presented digital twin framework is capable of enhancing the coordination of mineral processing efforts at all stages in the mine life cycle. Moreover, it has been shown as an important risk assessment tool in the management of decision-making processes.

The three components that were outlined in Section 2 represent three converging areas of knowledge that constitute our methodology

- Machine-Learning
- Discrete Event Simulation
- Digital Twin Development

The sample computations that were presented in Section 3 are indeed a result of this methodology (Figure 1), exploring the functionality of a particular random forest model. Moreover, the same approach can be used for developing and evaluating other potential strategies toward machine-learning-enabled process control. These potential alternatives may explore different types of machine-learning models, but even keeping with the same model, the methodology can explore different approaches to data acquisition.

Vein-hosted gold deposits in particular, may be subject to decreasing grades, mineralogical changes, and other orebody complexities. The present case study highlights the importance of integrated management decision tools to assess the effects of geological uncertainty on system performance. In this example, the importance of the target total ore stockpile as a control variable, to act as a buffer and stabilize plant feed, is aligned with previous results by Órdenes et al. [31]. Furthermore, this study provides an alternative solution to a problem that may arise in many brownfield mining projects that are seeking to expand reserves and ultimately mine life.

The continued deposit characterization and exploration for new resources are important activities for any mining project, and result in the continuous addition of new information that can impact potential resource evaluation and downstream processing options. It is, therefore, critical that the digital twin be routinely updated as new sample data are acquired through both infill and exploratory drill programs. The current version of the framework was presented for offline applications in the design stages of a mining project, which would require the manual transfer of data between physical and virtual systems. However, the intent would be to eventually integrate the digital twin into a mine control system with automatic bidirectional data exchange as development progresses and operational data accumulates. This gradual transition is key to ensuring the digital framework is both accurate and valid for its specific application prior to full implementation.

Future work will include a comprehensive representation of positional information (i.e., sample coordinate data) for better simulation of the orebody; this will enable the coordination of drill core sampling, leading to further advances in control and digital twin development. There is particular interest in developing the techniques for low grade deposits since these could be clear example of value generation, if it results in an expansion of the definable economic reserves. Moreover, part of the research will be to adapt developments that are occurring outside of the mining industry. As an example, the IoTwins Project has developed a multi-layered computing architecture that empowers small and medium enterprises [45] and its conceptualization of the Internet of Things (IoT) considers a distributed and mobile infrastructure that may support automated hauling for underground mining [43]. Our advances in simulation methodologies will gain more relevance through collaborations that include communications service providers and other technology developers [59,60].

Author Contributions: Conceptualization, F.P.-G., J.Ó. and A.N.; methodology, F.P.-G. and J.Ó.; data curation, F.P.-G.; investigation, J.Ó. and F.P.-G.; writing—original draft preparation, F.P.-G. and J.Ó.; writing—reviewing and editing, R.W. and A.N.; supervision, A.N.; funding acquisition, A.N. All authors have read and agreed to the published version of the manuscript.

Funding: Funding for this work was provided by NSERC, grant No. 2020-04605, supported by the Canadian government. This research was partially funded by the National Agency for Research and Development of Chile—ANID (J.Ó.). A research stipend was also received through the Mitacs Accelerate program in partnership with Watts, Griffis and McOuat Limited (J.Ó.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the staff at the Florida Mine (Alhué, Chile) for their assistance and support in conceptualizing and describing their operational problems and providing preliminary data. Thanks are due in particular to Gonzalo Lara, Carlos Montalvo, and César Aguilera.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gorain, B.K.; Kondos, P.D.; Lakshmanan, V.I. Innovations in Gold and Silver Processing. In *Innovative Process Development in Metallurgical Industry*; Springer International Publishing: Cham, Switzerland, 2016; pp. 393–428.
2. Navarra, A.; Grammatikopoulos, T.; Waters, K. Incorporation of Geometallurgical Modelling into Long-Term Production Planning. *Miner. Eng.* **2018**, *120*, 118–126. [\[CrossRef\]](#)
3. Dominy, S.C.; Stephenson, P.R.; Annels, A.E. Classification and Reporting of Mineral Resources for High-Nugget Effect Gold Vein Deposits. *Explor. Min. Geol.* **2001**, *10*, 215–233. [\[CrossRef\]](#)
4. Morales, N.; Seguel, S.; Cáceres, A.; Jélvez, E.; Alarcón, M. Incorporation of Geometallurgical Attributes and Geological Uncertainty into Long-Term Open-Pit Mine Planning. *Minerals* **2019**, *9*, 108. [\[CrossRef\]](#)
5. Darling, P. *SME Mining Engineering Handbook*, 3rd ed.; Society for Mining, Metallurgy, and Exploration, Inc.: Englewood, NJ, USA, 2011; ISBN 978-0-87335-341-0.
6. Callaway, G.; Ramsbottom, O. Can the Gold Industry Return to the Golden Age? McKinsey & Company. 2019. Available online: <https://www.mckinsey.com/~{} /media/mckinsey/industries/metals%20and%20mining/our%20insights/can%20the%20gold%20industry%20return%20to%20the%20golden%20age/can-the-gold-industry-return-to-the-golden-age-vf.pdf> (accessed on 23 October 2021).
7. Burkov, A. *The Hundred-Page Machine Learning Book*; Andriy Burkov: Quebec City, QC, Canada, 2019; ISBN 978-1-9995795-1-7.
8. Ali, D.; Frimpong, S. Artificial Intelligence, Machine Learning and Process Automation: Existing Knowledge Frontier and Way Forward for Mining Sector. *Artif. Intell. Rev.* **2020**, *53*, 6025–6042. [\[CrossRef\]](#)
9. McCoy, J.T.; Auret, L. Machine Learning Applications in Minerals Processing: A Review. *Miner. Eng.* **2019**, *132*, 95–109. [\[CrossRef\]](#)
10. Cisternas, L.A.; Lucay, F.A.; Botero, Y.L. Trends in Modeling, Design, and Optimization of Multiphase Systems in Minerals Processing. *Minerals* **2020**, *10*, 22. [\[CrossRef\]](#)
11. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
12. Breiman, L. Bagging Predictors. *Mach. Learn.* **1996**, *24*, 123–140. [\[CrossRef\]](#)
13. Nembrini, S.; König, I.R.; Wright, M.N. The Revival of the Gini Importance? *Bioinformatics* **2018**, *34*, 3711–3718. [\[CrossRef\]](#)
14. Guo, K.; Wan, X.; Liu, L.; Gao, Z.; Yang, M. Fault Diagnosis of Intelligent Production Line Based on Digital Twin and Improved Random Forest. *Appl. Sci.* **2021**, *11*, 7733. [\[CrossRef\]](#)
15. Belgiu, M.; Drăgu, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [\[CrossRef\]](#)
16. Oshiro, T.M.; Perez, P.S.; Baranauskas, J.A. How Many Trees in a Random Forest? In *Machine Learning and Data Mining in Pattern Recognition*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7376, pp. 154–168. [\[CrossRef\]](#)
17. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random Forests for Land Cover Classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [\[CrossRef\]](#)
18. Fawcett, T. An Introduction to ROC Analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [\[CrossRef\]](#)
19. Cassandra, C.G.; Lafortune, S. *Introduction to Discrete Event Systems*, 2nd ed.; Cassandra, C.G., Lafortune, S., Eds.; Springer: Boston, MA, USA, 2008; Volume 37, ISBN 978-0-387-33332-8.
20. Upadhyay, S.P.; Askari-Nasab, H.; Tabesh, M.; Badiozamani, M.M. Simulation and Optimization in Open Pit Mining. In *Proceedings of the Application of Computers and Operations Research in the Mineral Industry—Proceedings of the 37th International Symposium APCOM 2015*, Fairbanks, AK, USA, 23–25 May 2015; Society for Mining, Metallurgy & Exploration, Inc.: Englewood, NJ, USA, 2015; pp. 532–543.
21. Wilson, R.; Mercier, P.; Patarachao, B.; Navarra, A. Partial Least Squares Regression of Oil Sands Processing Variables within Discrete Event Simulation Digital Twin. *Minerals* **2021**, *11*, 689. [\[CrossRef\]](#)
22. Peña-Graf, F.A.; Grammatikopoulos, T.; Kabemba, A.; Navarra, A. Integrated Feed Management of Mineral Processing Plants with Application to Chromite Processing. *Can. Metall. Q.* **2021**, *60*, 130–136. [\[CrossRef\]](#)

23. Fahl, S.K. *Benefits of Discrete Event Simulation in Modeling Mining Processes*; University of Alberta: Edmonton, AB, Canada, 2017.
24. Wilson, R.; Toro, N.; Naranjo, O.; Emery, X.; Navarra, A. Integration of Geostatistical Modeling into Discrete Event Simulation for Development of Tailings Dam Retreatment Applications. *Miner. Eng.* **2021**, *164*, 106814. [\[CrossRef\]](#)
25. Jung, D.; Baek, J.; Choi, Y. Stochastic Predictions of Ore Production in an Underground Limestone Mine Using Different Probability Density Functions: A Comparative Study Using Big Data from ICT System. *Appl. Sci.* **2021**, *11*, 4301. [\[CrossRef\]](#)
26. Navarra, A.; Rafiei, A.A.; Waters, K. A Systems Approach to Mineral Processing Based on Mathematical Programming. *Can. Metall. Q.* **2017**, *56*, 35–44. [\[CrossRef\]](#)
27. Navarra, A.; Wilson, R.; Parra, R.; Toro, N.; Ross, A.; Nave, J.-C.; Mackey, P.J. Quantitative Methods to Support Data Acquisition Modernization within Copper Smelters. *Processes* **2020**, *8*, 1478. [\[CrossRef\]](#)
28. Navarra, A.; Alvarez, M.; Rojas, K.; Menzies, A.; Pax, R.; Waters, K. Concentrator Operational Modes in Response to Geological Variation. *Miner. Eng.* **2019**, *134*, 356–364. [\[CrossRef\]](#)
29. Saldaña, M.; Neira, P.; Flores, V.; Moraga, C.; Robles, P.; Salazar, I. Analysis of the Dynamics of Rougher Cells on the Basis of Phenomenological Models and Discrete Event Simulation Framework. *Metals* **2021**, *11*, 1454. [\[CrossRef\]](#)
30. Saldaña, M.; Toro, N.; Castillo, J.; Hernández, P.; Navarra, A. Optimization of the Heap Leaching Process through Changes in Modes of Operation and Discrete Event Simulation. *Minerals* **2019**, *9*, 421. [\[CrossRef\]](#)
31. Órdenes, J.; Wilson, R.; Peña-Graf, F.; Navarra, A. Incorporation of Geometallurgical Input into Gold Mining System Simulation to Control Cyanide Consumption. *Minerals* **2021**, *11*, 1023. [\[CrossRef\]](#)
32. Bergmann, S.; Feldkamp, N.; Strassburger, S. Emulation of Control Strategies through Machine Learning in Manufacturing Simulations. *J. Simul.* **2017**, *11*, 38–50. [\[CrossRef\]](#)
33. Glowacka, K.J.; Henry, R.M.; May, J.H. A Hybrid Data Mining/Simulation Approach for Modelling Outpatient No-Shows in Clinic Scheduling. *J. Oper. Res. Soc.* **2009**, *60*, 1056–1068. [\[CrossRef\]](#)
34. Greasley, A. Architectures for Combining Discrete-Event Simulation and Machine Learning. In Proceedings of the 10th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, SIMULTECH 2020, Paris, France, 8–10 July 2020; pp. 47–58. [\[CrossRef\]](#)
35. Greasley, A.; Edwards, J.S. Enhancing Discrete-Event Simulation with Big Data Analytics: A Review. *J. Oper. Res. Soc.* **2021**, *72*, 247–267. [\[CrossRef\]](#)
36. Singh, M.; Fuenmayor, E.; Hinchy, E.P.; Qiao, Y.; Murray, N.; Devine, D. Digital Twin: Origin to Future. *Appl. Syst. Innov.* **2021**, *4*, 36. [\[CrossRef\]](#)
37. Grieves, M.; Vickers, J. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems (Excerpt). Available online: https://www.researchgate.net/publication/307509727_Origins_of_the_Digital_Twin_Concept (accessed on 23 October 2021).
38. Shafto, M.; Rich, M.C.; Glaessgen, D.E.; Kemp, C.; Lemoigne, J.; Wang, L. *Modeling, Simulation, Information Technology & Processing Roadmap. National Aeronautics and Space Administration*; NASA Headquarters: Washington, DC, USA, 2010; Volume 32, pp. 1–38.
39. IBM. What Is a Digital Twin? Available online: <https://www.ibm.com/topics/what-is-a-digital-twin> (accessed on 23 October 2021).
40. Mehra, A. Digital Twin Market Worth \$48.2 Billion by 2026. Available online: <https://www.marketsandmarkets.com/PressReleases/digital-twin.asp> (accessed on 23 October 2021).
41. Moore, E. *CIM Magazine*; Canadian Institute of Mining, Metallurgy and Petroleum: Westmount, QC, Canada, 2018.
42. Carpenter, J.; Cowie, S.; Stewart, P.; Jones, E. *Offer A Machine Learning at a Gold-Silver Mine: A Case Study from the Ban Houayxai Gold-Silver Operation*; The Australasian Institute of Mining and Metallurgy: Parkville, Australia, 2018.
43. Underground Communications Infrastructure Sub-Committee of the Underground Mining Working Group. *Underground Mine Communications Infrastructure Guidelines Part III: General Guidelines*; Global Mining Guidelines Group: Ormstown, QC, Canada, 2019.
44. Hargrave, C.O.; Ralston, J.C.; Hainsworth, D.W. Optimizing Wireless LAN for Longwall Coal Mine Automation. *IEEE Trans. Ind. Appl.* **2007**, *43*, 111–117. [\[CrossRef\]](#)
45. Costantini, A.; Duma, D.C.; Martelli, B.; Antonacci, M.; Galletti, M.; Tisbeni, S.R.; Bellavista, P.; di Modica, G.; Nehls, D.; Ahouangonou, J.-C.; et al. A Cloud-Edge Orchestration Platform for the Innovative Industrial Scenarios of the IoTwins Project. In *Computational Science and Its Applications—ICCSA 2021*; Springer: Cham, Switzerland, 2021; Volume 12950, pp. 533–543.
46. Lafrance, B. Structural Controls on Hydrothermal Lode Gold Deposits. In *Proceedings of the Manitoba Mining & Minerals Convention*; Government of Manitoba: Winnipeg, MB, Canada, 2008.
47. Hodgson, C.J. The Structure of Shear-Related, Vein-Type Gold Deposits: A Review. *Ore Geol. Rev.* **1989**, *4*, 231–273. [\[CrossRef\]](#)
48. Adler, L.; Thompson, S.D. Mining Methods Classification System. In *SME Mining Engineering Handbook*; Society for Mining, Metallurgy, and Exploration, Inc.: Englewood, NJ, USA, 2011; pp. 349–357.
49. Cabello, J. Gold Deposits in Chile. *Andean Geol.* **2021**, *48*, 1–23. [\[CrossRef\]](#)
50. Jannas, R.; Araneda, R. Geología de la veta indio sur 3.500; una estructura tipo bonanza del yacimiento el indio. *Andean Geol.* **1985**, *24*, 49–62.
51. Jannas, R.R.; Bowers, T.S.; Petersen, U.; Beane, R.E. High-Sulfidation Deposit Types in the El Indio District, Chile. In *Geology and Ore Deposits of the Central Andes*; Society of Economic Geologists, Inc.: Littleton, CO, USA, 1999.

52. Sillitoe, R.H. Styles of High-Sulphidation Gold, Silver and Copper Mineralisation in Porphyry and Epithermal Environments. In *Proceedings of the Australasian Institute of Mining and Metallurgy*; Vic.: The Institute: Parkville, Australia, 2000; Volume 305, pp. 19–34.
53. Thompson, J.F.H.; Gale, V.G.; Tosdal, R.M.; Wright, W.A. Characteristics and Formation of the Jeronimo Carbonate-Replacement Gold Deposit, Potrerillos District, Chile. *Andean Metall. New Discov. Concepts Updates Spec. Publ.* **2004**, *11*, 75–95.
54. Lazcano, A.; Fuentes, H.M. Jeronimo, Un Nuevo Depósito Aurífero En El Área El Hueso—Agua de La Falda. In *Proceedings of the VIII Congreso Geológico Chileno*, Antofagasta, Chile, 13–17 October 1997; Universidad Católica del Norte: Antofagasta, Chile, 1997; pp. 1033–1037.
55. Órdenes, J. *Influencia de La Mineralogía de La Veta Bonanza En El Proceso Hidrometalúrgico de Extracción de Au y Ag, Yacimiento El Peñón, Chile*; Universidad Católica del Norte: Antofagasta, Chile, 2014.
56. Cetin, M.C.; Emre Altun, N.; Umit Atalay, M.; Buyuktanir, K. Bottle Roll Testing for Cyanidation of Gold Ores: Problems Related to Standardized Procedures on Difficult-to-Process Ores. In *Proceedings of the 3rd World Congress on Mechanical, Chemical, and Material Engineering (MCM'17)*, Rome, Italy, 8–10 June 2017. [[CrossRef](#)]
57. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
58. Wills, B.A.; Finch, J.A. Mass Balancing Methods. In *Wills' Mineral Processing Technology*; Butterworth-Heinemann: Oxford, UK, 2016; pp. 69–83.
59. Farrelly, C.T.; Davies, J. Interoperability, Integration, and Digital Twins for Mining—Part 2: Pathways to the Network-Centric Mine. *IEEE Ind. Electron. Mag.* **2021**, *15*, 22–31. [[CrossRef](#)]
60. Servin, M.; Vesterlund, F.; Wallin, E. Digital Twins with Distributed Particle Simulation for Mine-to-Mill Material Tracking. *Minerals* **2021**, *11*, 524. [[CrossRef](#)]