

Article

A Stochastic Model Approach for Modeling SAG Mill Production and Power Through Bayesian Networks: A Case Study of the Chilean Copper Mining Industry

Manuel Saldana ^{1,2,*} , Edelmira Gálvez ^{3,*} , Mauricio Sales-Cruz ⁴ , Eleazar Salinas-Rodríguez ⁵ , Jonathan Castillo ⁶ , Alessandro Navarra ⁷ , Norman Toro ¹ , Dayana Arias ⁸ and Luis A. Cisternas ² 

¹ Faculty of Engineering and Architecture, Universidad Arturo Prat, Iquique 1110939, Chile; notoro@unap.cl

² Departamento de Ingeniería Química y Procesos de Minerales, Universidad de Antofagasta, Antofagasta 1271155, Chile; luis.cisternas@uantof.cl

³ Departamento de Ingeniería Metalúrgica y Minas, Universidad Católica del Norte, Antofagasta 1270709, Chile

⁴ Departamento de Procesos y Tecnología, Universidad Autónoma Metropolitana—Cuajimalpa, Mexico City 05348, Mexico; asales@cua.uam.mx

⁵ Academic Area of Earth Sciences and Materials, Institute of Basic Sciences and Engineering, Autonomous University of the State of Hidalgo, Pachuca 42184, Mexico; salinasr@uaeh.edu.mx

⁶ Departamento de Ingeniería en Metalurgia, Universidad de Atacama, Copiapó 1531772, Chile; jonathan.castillo@uda.cl

⁷ Department of Mining and Materials Engineering, McGill University, 3610 University Street, Montreal, QC H3A 0C5, Canada; alessandro.navarra@mcgill.ca

⁸ Laboratory of Molecular Biology and Applied Microbiology, Centro de Investigación en Fisiología y Medicina de Altura (FIMEDALT), Departamento Biomédico, Facultad de Ciencias de la Salud, Universidad de Antofagasta, Antofagasta 1240000, Chile; dayana.arias@uantof.cl

* Correspondence: masaldana@unap.cl (M.S.); egalvez@ucn.cl (E.G.); Tel.: +56-953834174 (M.S.)

Abstract

Semi-autogenous (SAG) milling represents one of the most energy-intensive and variable stages of copper mineral processing. Traditional deterministic models often fail to capture the nonlinear dependencies and uncertainty inherent in industrial operations such as granulometry, solids percentage in the feeding or hardness. This work develops and validates a stochastic model based on Discrete Bayesian networks (BNs) to represent the causal relationships governing SAG Production and SAG Power under uncertainty or partial knowledge of explanatory variables. Discretization is adopted for methodological reasons as well as for operational relevance, since SAG plant decisions are typically made using threshold-based categories. Using operational data from a Chilean mining operation, the model fitted integrates expert-guided structure learning (Hill-Climbing with BDeu/BIC scores) and Bayesian parameter estimation with Dirichlet priors. Although validation indicators show high predictive performance ($R^2 \approx 0.85$ – 0.90 , $RMSE < 0.5$ bin, and micro-AUC ≈ 0.98), the primary purpose of the BN is not exact regression but explainable causal inference and probabilistic scenario evaluation. Sensitivity analysis identified water feed and solids percentage as key drivers of throughput (SAG Production), while rotational speed and pressure governed SAG Power behavior. The BN framework effectively balances accuracy and interpretability, offering an explainable probabilistic representation of SAG dynamics. These results demonstrate the potential of stochastic modeling to enhance process control and support uncertainty-aware decision making.

Keywords: Bayesian networks; stochastic process modelling; SAG milling; comminution; mineral processing



Academic Editors: Sunil Kumar Tripathy, Javier Ruiz del Solar, Humberto Estay and Pia Lois-Morales

Received: 9 November 2025

Revised: 23 December 2025

Accepted: 30 December 2025

Published: 6 January 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Copper mining is a growing industry [1], and of the copper minerals found on the planet, the majority are sulfides and a smaller amount are oxides. Therefore, flotation techniques and smelting processes are used to process these minerals, and to a lesser extent, hydrometallurgical techniques. Flotation techniques generate a large amount of waste, resulting in tailings and the generation of acid drainage due to the oxidation of minerals with a high presence of pyrite. On the other hand, while pyrometallurgical processes generate substantial amounts of sulfur dioxide (SO₂), which when combined with NO_x and CO₂ can cause significant problems such as acid rain and increased local pollution [2], hydrometallurgy offers a more ecological option for the processing of both oxidized and sulfide minerals [3].

For the last 50 years, heap leaching processes have been an attractive technological alternative for the processing of low-grade ores [4]. Hydrometallurgical processes are applied to minerals previously crushed in crushers, where the copper present in the mineralized rock is extracted by adding a solution composed of water and leaching agents such as H₂SO₄, among others [5] and can be used for low to medium-grade oxidized copper minerals (0.3%–0.7%), as well as for secondary copper sulfides. It is important to note that in heap leaching, particle sizes greater than 6 mm and less than 10–40 mm are preferred to maintain adequate heap permeability. Clay minerals can cause clogging over time due to swelling and gradual disintegration, so sizes smaller than 6 mm are recommended to be avoided. The leaching solution is applied to the top of the heap by sprinklers, and gravity causes it to flow down through the heap [6].

Then, due to the progressive depletion of oxidized copper deposits in Chile, mining has driven the move towards the exploitation of sulfide minerals in recent decades, which have become the main source of recoverable copper [7]. These sulfide minerals, which mainly include sulfides such as chalcopyrite, covellite and chalcocite, have higher concentrations of copper but require pyrometallurgical routes such as flotation concentration and smelting, since they are practically insoluble in conventional hydrometallurgical methods [8]. In the Chilean context, where around 80% of copper production originates from sulfide minerals, this transition reflects a significant technological change aimed at maintaining the economic viability of the industry in the face of oxide depletion [9].

Pyrometallurgical processes, unlike hydrometallurgy, are procedures that use high temperatures to process copper concentrates and produce refined copper or blister copper, and require greater resources to reduce the size of the ore destined to feed the flotation stages [10]. In the pyrometallurgical route, comminution processes (comprising crushing and grinding) are energy-intensive; specifically, 70% of the energy costs are allocated to ore comminution. Within the grinding section, the grinding circuit alone accounts for up to 97% of the energy cost. This highlights that grinding is an important area where considerable savings could be achieved in mineral processing circuits [11]. In this context, implementing technological improvements in energy efficiency, such as the use of more efficient equipment, optimization of grinding circuits, and application of alternative energies, not only reduces consumption and costs, but also contributes significantly to more sustainable and environmentally responsible mining [12,13].

Within comminution, semi-autogenous grinding (SAG) is a critical stage in the copper value chain, as it receives large crushed ore and prepares it for subsequent concentration stages. This process represents one of the largest energy consumers in mining, with a significant impact on operational costs and overall plant productivity [14]. However, the operation of SAG mills is subject to high variability, derived from factors such as ore hardness, feed particle size distribution, fill level, and process water flow rate. This

variability introduces uncertainty into production prediction and makes it difficult to identify optimal operating conditions [15].

Traditionally, SAG mill performance has been modeled using deterministic or empirical approaches such as energy balances, power equations, or simulations on platforms like JKSimMet [16]. While these methods have provided useful results under controlled conditions, they present significant limitations when faced with the uncertainty inherent in mining processes, characterized by high variability in ore hardness, feed grain size distribution, operational conditions, or other geo-metallurgical variables. Such models typically assume linear or fixed relationships between variables [17], which restricts their ability to represent complex conditional dependencies or adapt to scenarios with incomplete information [18–21]. Then, the generation of models that allow incorporating variables such as the uncertainty inherent in the variability of the mineral and operational conditions, representing non-linear conditional relationships between multiple variables and combining both historical data and expert knowledge, represents a superlative advance in the development of more robust and flexible models, capable of predicting scenarios under incomplete information and supporting decision making aimed at improving process efficiency and operational sustainability.

In this study, a stochastic approach based on Bayesian networks (BNs) [22] is proposed to model the SAG milling process. This method considers both the distributions of the independent variables and the conditional dependencies among them, enabling a probabilistic representation of how changes in operational conditions influence process responses. Bayesian modeling is robust to incomplete evidence or partial knowledge of the system, allowing the network to produce meaningful probabilistic inferences even under uncertainty in the explanatory variables. In this context, discretization of continuous variables is adopted for methodological and operational reasons. From a methodological standpoint, discretization improves the stability of structure learning in Bayesian networks, reduces sensitivity to outliers, and avoids imposing restrictive linear or Gaussian assumptions that may not hold in highly variable SAG mill environments. From an operational perspective, SAG plant decisions are typically made using threshold-based categories (e.g., low, normal, or high production and power levels), making discretized representations more aligned with real decision-making practices. Accordingly, the primary purpose of this modeling framework is not exact numerical regression, but rather explainable causal inference and probabilistic scenario evaluation.

Therefore, the objective of this study is to develop and validate a stochastic model based on BN to characterize the Production and Power behavior of a SAG mill, considering operational uncertainty and conditional dependencies between variables. This approach seeks to contribute to a better understanding of the factors that determine process efficiency and support decision making aimed at optimizing operations. It is worth noting that the present study focuses on a static BN formulation; dynamic extensions required for a full probabilistic digital twin, including temporal dependencies and online updating under non-stationary conditions, are identified as future work.

2. Background

2.1. Grinding Modeling

Modeling approaches to SAG mill performance in copper mining are diverse and seek to optimize the grinding process, improve energy efficiency, and optimize operational control. These models range from dynamic state-space models to advanced control systems and hybrid algorithms. Each approach offers unique insights and solutions to the challenges facing SAG mill operations, particularly in the context of copper mining.

The SAG grinding process has been studied and modeled by several authors, generating explanatory models of grinding, either as an individual process or integrated into aggregate processes, such as the mine-to-mill (M2M) paradigm, a practice that has allowed analytically evaluating aspects of mining and processing, as well as executing models and simulations to predict the effects of mine variations on downstream processing [23]. It is possible to identify different trends in the modeling, design, simulation and optimization of complex systems such as mineral processing, using techniques such as computational fluid dynamics (CFD), response surface methodology (RSM), machine learning (ML) algorithms such as artificial neural networks (ANNs), support vector machines (SVM) or random forest (RF), in addition to uncertainty analysis (UA) or sensitivity analysis (SA) [24]. Next, various theoretical models are presented, including those expressed through mathematical relationships or equations, as well as models that are based on ML.

As usual, mill power equations are derived from mechanics and are used to predict the power output of ball/AG/SAG mills [25–27]. However, few models consider the feed size distribution as a design variable. Numerous works using the Mine-to-Mill (M2M) approach [28–30] have demonstrated the influence of the mineral ROM (in addition to characteristics such as hardness, granulometry or lithology, among others) on the comminution process performance [29], where the feed size distribution is even more influential (in many cases) than the mineral characteristics themselves [28,31]. Furthermore, there is a relationship between the feed mineral, the mill geometry and the operating conditions with the specific energy of the circuit [32]. In this same M2M line, the feed size distribution directly affects the SAG mill load, which in turn affects the mill power and impacts its efficiency [33].

Silva et al. [34] fitted specific power and energy equations for the design of SAG mills, using them to predict energy consumption as a function of mill size, internal load and density level, critical speed and feed size distribution. Dong et al. [35] analyzed the factors influencing energy consumption in grinding processes, establishing three models for energy prediction: regression models, ANN and a hybrid genetic algorithm (GA)–ANN model. Additionally, Lucay et al. [36,37] showed that UA and global sensitivity analysis (GSA) are useful tools to identify operational conditions under uncertainty in grinding systems, while Li et al. [38] evaluate the impact of active speed control and closed-circuit adjustment of the cone crusher on the performance and/or energy efficiency of the SAG circuit.

Asghari et al. [39] study the effects of ore characteristics and operational parameters on mill performance using mathematical models such as t_{10} or Bond work index, as well as mill product particle shape to monitor fracture events as a function of ore strength. Lvov et al. [40] and Marijnissen et al. [41] develop SAG mill models based on the discrete element method and computational fluid dynamics (CFD), which allows the determination of the energy–roughness relationships of the grinding process (SAG) under certain conditions [40], or obtaining the velocities and collision angles of a representative group of particles in the mill [41].

On the other hand, numerous review articles have discussed [42,43] and applied [44] the potential need for process engineers to leverage tools such as applied mathematics, statistics, ML, and AI. From the literature review, highlighted applications are the analysis of mill load measurement mechanisms and methods based on mechanical vibrations and acoustic signals [45]; the impact of blast fragmentation control on increasing mill throughput [46]; the development of a dynamic model of a SAG mill using equations based on the conventional non-stationary population balance approach [47]; the identification of the best operating conditions (e.g., cut size) for optimal grinding using a gradient recovery model [48]; case studies of grinding circuit modeling by tuning support vector machine

algorithms [49], in order to analyze descriptor variables such as power or temperature [50]; inferential measurement of SAG mill parameters [51,52]; multicomponent phenomenological modeling, which represents the performance of a SAG mill as a function of mineral feed distribution and components [53]; and energy consumption prediction modeling [15,54], among others.

Additional applications of machine learning (ML) to SAG mill modeling focus on circuit control, predicting energy consumption [15,54]. In Avalos et al. [15], ML-based predictive methods (regressions, kNN, SVM and ANN) were studied for estimating SAG mill energy consumption (based on variables such as feed tonnage, mill pressure and mill speed). In Kahraman et al. [54] a data-driven multi-rate (MRA) method was developed to predict SAG mill energy consumption, using deep learning as the prediction model for the MRA method.

In Olivier et al. [55], decision trees were used to model decision making regarding the removal of critical-sized material from the circuit to avoid mill overloading, and to extract rules to guide operators in making specific decisions. Azizi et al. [56] investigated the application of supervised algorithms (single- and multi-kernel SVM regression analysis and ANN) to model grinding media consumption rates based on multiple input factors, concluding that multi-kernel SVM can be efficiently used to model such consumption rates. In Hoseinian et al. [57] a hybrid model combining ANN and genetic algorithms (GANNs) was used to predict SAG mill power. This model considered various operating parameters such as feed moisture and mill load and has proven to be more efficient than traditional ANN models.

Other SAG mill modeling methods include adaptive dynamic programming, discrete element methods (DEMs), and fuzzy logic and expert systems. Adaptive dynamic programming and reference controllers have been proposed for the operational control of mineral milling processes. These systems focus on maintaining the optimal particle size and circulating load, using data-driven methods to manage input constraints and optimize control without detailed system modeling [58]. DEM has been used to simulate the particle breakage process in a SAG mill, focusing on parameters such as mill speed ratio, fill level ratio, and steel ball ratio. This method helps understand grinding resistance and energy consumption by providing insight into optimal operating conditions [59]. Supervisory fuzzy expert controllers have been developed for SAG mill circuits, which use the Mamdani method to calculate optimal setpoints for the plant control loops [60]. This system is designed to maintain full-capacity operations and is integrated with plant-level control systems.

The literature reveals a clear evolution in SAG mill modeling, from simple empirical equations to sophisticated ML approaches. However, significant shortcomings remain in the development of models that can simultaneously manage the uncertainty inherent in industrial operations, capture conditional dependencies between variables, and provide interpretable results for operational decision making. Current ML approaches, while achieving high prediction accuracy, often lack the transparency necessary for industrial implementation and fail to quantify prediction uncertainty. Classical mechanistic models, although interpretable, cannot adequately capture the complex, nonlinear relationships and stochastic variability present in real-life SAG mill operations.

BNs represent a promising solution to these limitations, as they offer a probabilistic modeling framework that can integrate prior knowledge with data-driven learning, seamlessly manage incomplete information, and provide uncertainty quantification while maintaining model interpretability. The successful application of Bayesian methods in related domains, coupled with the identified shortcomings in current SAG mill modeling approaches, strongly justify the need to investigate the applications of Bayesian

networks for robust and interpretable prediction of SAG mill performance. This will not only allow for more reliable predictive and decision-support tools aimed at improving efficiency, but also has the potential to reduce energy consumption and move towards more sustainable mining.

2.2. Handling Uncertainty and Conditional Dependencies

Recent work addresses the uncertainty, probabilistic relationships, and computational challenges that arise when incorporating uncertainty into process models, linking the methods to the mining context and inference objectives. The authors recommend explicit probabilistic models (BN, hierarchical Bayesian, Gaussian process-based probabilistic controllers) and their hybridization with simulators to capture both aleatory and epistemic uncertainty. Some substantive points and evidence are presented below:

- BN for Conditional Structure and Uncertainty: Bayesian graphical models explicitly encode conditional dependencies and provide probabilistic predictions and sensitivity paths that help explain the causes of events in grinding circuits, as shown in the SAG BN case study, where the fresh ore feed rate is modeled [61].
- Hierarchical Bayesian Formulations and Event/Measurement Uncertainty: Recent hierarchical Bayesian applications to SAG performance incorporate nested structure (e.g., ore-level hardness, operational set values) to improve uncertainty quantification and resilience to different conditions [62]. Additionally, the process mining literature emphasizes that many industrial event records suffer from time-stamp/value inaccuracy and that representing uncertainty increases model expressiveness but also computational complexity. Graph-based approximation or representations can mitigate combinatorial explosion [63].
- Practical guidance and challenges: Reviews of Bayesian methods for mineral processing and broader studies note that Bayesian approaches allow for principled uncertainty quantification and model comparison, but require careful prior/domain integration and reliable sensor data [64].
- Probabilistic methods at the controller level: Probabilistic model predictive control using Gaussian processes has been proposed for flotation circuits to propagate uncertainty through control decisions, illustrating viable probabilistic control paths for metallurgical units [65].

2.3. Model Comparisons, Hybrid Dynamics, Digital Twins and Optimization

Table 1 presents a concise comparison of different models (Bayesian networks—BNs, artificial neural networks—ANNs, support vector machines—SVM, random forests/tree ensembles—RF), highlighting their advantages, data requirements, uncertainty management, and results of previous studies.

Then, some key comparative notes are presented below:

- Prediction versus explanation: BNs sacrifice some predictive flexibility in exchange for explicit conditional structure and greater probabilistic interpretability, making it easier to identify underlying causes and make decisions [61,62].
- Temporal/sequence data: In power prediction in SAG circuits, recurrent architectures outperformed static regression models in an industrial study [15].
- Model ensembles and optimization: Tree ensembles (Cat-Boost/XGBoost/Random Forest) typically achieve high predictive accuracy on large industrial datasets and are easily integrated with parameter optimization algorithms [67,68].

Table 1. Comparative analysis of the models according to their main characteristics.

Model	Strengths	Typical Data Needs	Uncertainty Handling	Example Source
BN	Captures conditional dependencies and causal structure; interpretable probabilistic outputs	Moderate (structure + conditional tables); benefits from expert priors	Native probabilistic inference and sensitivity analysis [61,62]	Videla & Vargas SAG BN for feed prediction [61]
ANN (MLP; RNN; LSTM)	Strong nonlinear function approximation; RNNs excel for temporal forecasts	High for generalization; can be reduced with scenario-based design [15,66]	Usually point forecasts; uncertainty via Bayesian NN or ensemble methods	Avalos et al. found RNNs best for SAG energy forecasting [15]; Moraga et al. achieved $R^2 > 0.99$ using discretized scenarios [66]
SVM	Effective in moderate-dimensional regression/classification; robust regularization	Moderate	Non-probabilistic by default; probability estimates via calibration	Used among candidate predictors for SAG energy; often outperformed by ensembles in some tasks [15,67]
RF	Strong off-the-shelf performance, robust to outliers, interpretable feature importance	Moderate; tolerates noisy/heterogeneous inputs	Can provide empirical uncertainty (quantile forests/ensembles)	RFR outperformed SVR for filter cake moisture [67]; CatBoost was top predictor in a hybrid SAG throughput study [68]

Another paradigm that has gained traction in recent decades is the combination of probabilistic/Bayesian methods with simulation, discrete event simulation (DES), digital twins, or optimization frameworks to create dynamic or predictive decision systems; the first sentences explain the rationale for the hybridization (uncertainty + physics + data), while some key points and examples are presented below:

- Hybridization of digital twins and data-driven models: Several reviews advocate the integration of instrument measurements, machine learning-based simulation models, and digital twins for process control and scenario analysis in the mineral beneficiation chain [69].
- System-level coupling: Mining-metallurgical and techno-economic assessments combine stochastic (long-term) mine planning with SED (medium-term) and predictive models to evaluate technological improvements in the face of geological uncertainty, demonstrating practical hybrid workflows for decision making [70].
- Simulation models + optimizers: Studies that combine high-accuracy machine learning predictors (tree ensembles or neural networks) with evolutionary optimizers convert simulation models into optimal parameters to maximize SAG plant performance within constraints [68,71].
- Probabilistic control loop: A study proposing probabilistic methods using Gaussian processes for flotation shows how predictive uncertainty estimates can be integrated into control constraints and cost/objective functions of metallurgical units [65].

3. Materials and Methods

3.1. Study Case

This case study focuses on the milling process of a copper concentrator plant located in the Antofagasta region of northern Chile, an open-pit mining operation that extracts sulfide copper ores from mineral deposits (see circuit in Figure 1). The plant's main grinding line consists of a SAG mill, which will be represented by fitting a stochastic model (Bayesian Network). The comminution circuit of the concentrator plant consists of a 12.2 m × 7.3 m SAG mill, followed by two parallel ball mills for secondary grinding. In addition, pebbles produced by the mill are fed back to the SAG mill.

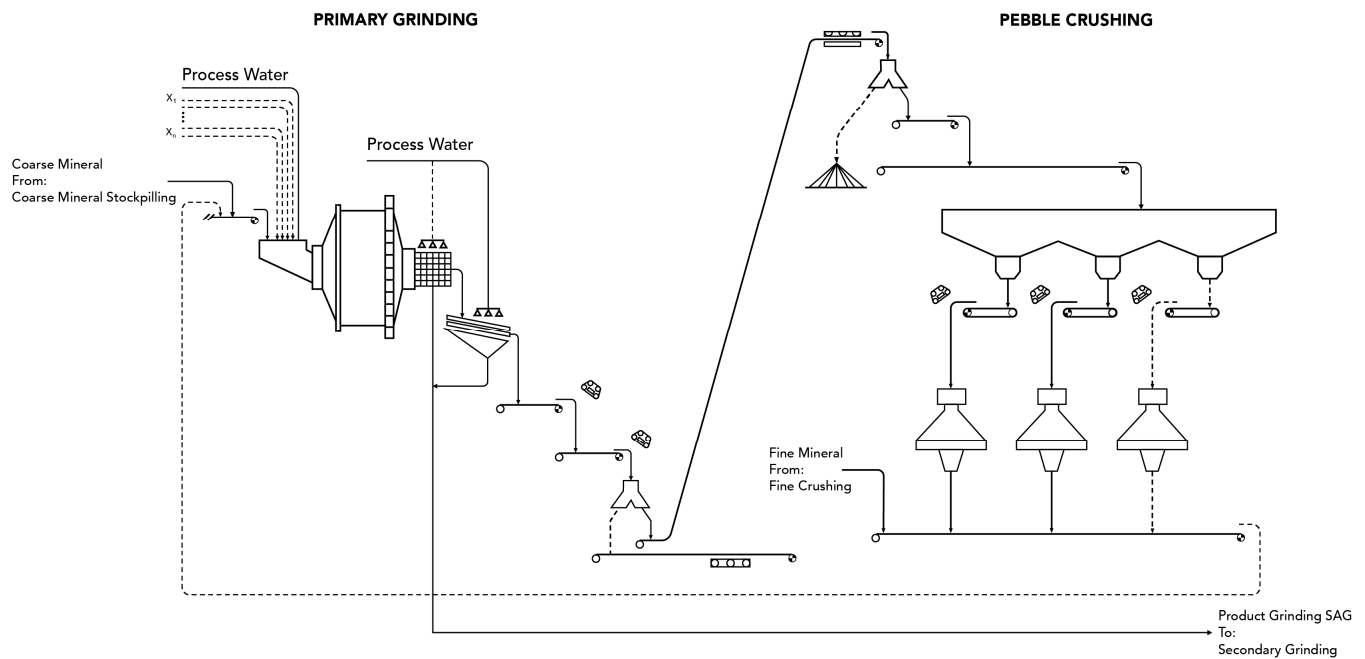


Figure 1. Primary grinding circuit layout.

Historical operating data was collected over a period of approximately 2 years, including 12 operational variables/parameters that are used as inputs to generate a representative model of the SAG mill. Independent variables considered were P_{80} , water flow rate, mill rotational speed, mill pressure, stockpile level, sump level, ore hardness, percentage of solids in the feed, pebbles, granulometry > 100 mm, granulometry < 30 mm, and liner age. The response variables correspond to the production rate in tons per hour (TpH) and the power consumption in megawatts (MW).

The variables/parameters considered for the modeling are described below.

- P_{80} : Size of the mesh opening that allows the passage of 80% of the granulometry.
- SAG water feeding (m^3/h): Water flow feeding to the SAG mill.
- SAG rotational speed (RPM): Mill rotational speed.
- SAG pressure (PSI): Mill fill or load level.
- Stockpile level (m): Stockpile level in the feeding stack.
- Sump level (m): Thicker downloading pool at the SAG mill.
- Hardness: Resistance offered by the mineral to abrasion or scraping.
- Solids in the feeding (%): Percentage of solids in the feed pulp.
- Pebbles (TpH): Pebbles (pebbles, chunks, or small stones) are the result of mineral grinding. These are hard materials and are difficult to reduce to a smaller size in the SAG mill.
- Granulometry > 100 mm (%): Percentage of the ore feed whose granulometry is greater than 100 mm.
- Granulometry < 30 mm (%): Percentage of the ore feed whose granulometry is less than 30 mm.
- Liner age (months): Categorical variable. Age of mill liners. Liners are part of the mill and act as protective sleeves for the internal casing (shell), which wears over time due to the strong and constant internal impact between the ore charge and the steel balls.

The proposed process for modeling SAG mill dynamics (see Figure 2) involves fitting a Bayesian network (after preprocessing the available data) to determine a model capable of representing responses based on detailed or partial knowledge of the feed variables. Once the network has been fitted, different scenarios can be simulated, evaluating the hypothesis

tests associated with the estimation of the explained variable, thus finding the optimal configurations, i.e., the variables (or their ranges) that maximize the production indicators.

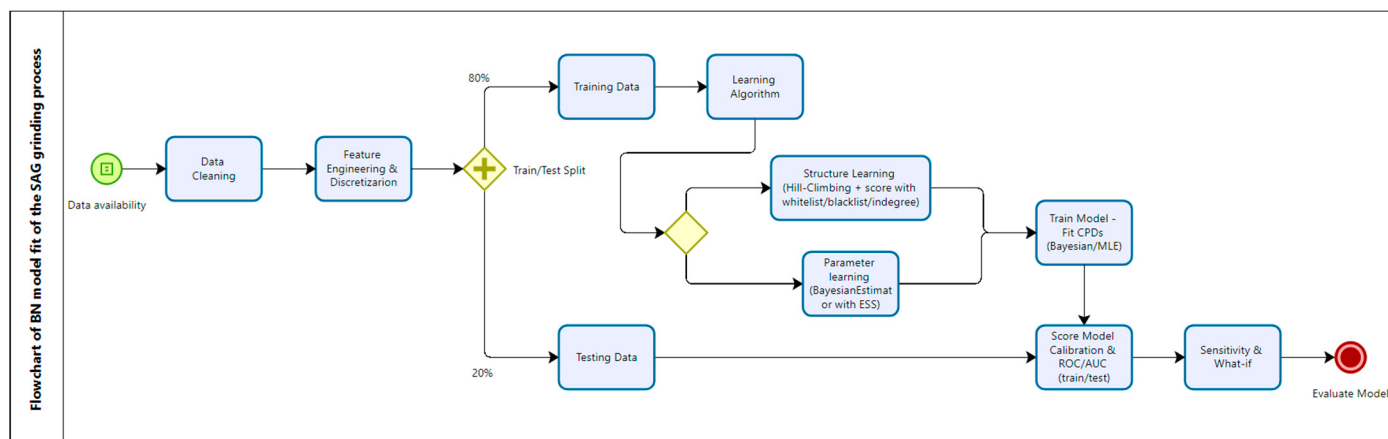


Figure 2. Flowchart of SAG milling process model fitting (developed in software Bizagi Modeler—version 4.2.0.003 [72]).

3.2. Machine Learning

ML techniques have an increasing presence and impact in a wide variety of research fields. McCoy and Auret [73] developed a review on the status of ML applications in mineral processing. Some applications include the prediction of grinding phase indicators, determining the chemical properties that have the greatest impact on grinding capacity indices by configuring ANNs [74–77], or regressions with SVM [78]. Modeling applications to predict mill performance indicators, based on process measurements, include the use of multivariate statistical methods such as Partial Least Squares (PLS) and Radial Basis Neural Networks (RBF), demonstrating that variables such as mill slurry density and ball charge volume can be reliably estimated from different operational characteristics [79]. In contrast, Ahmadzadeh and Lundberg [75] devised an approach to predict the remaining life of a mill re-liner while the mill is still in operation by constructing an artificial neural network capable of identifying complex connections between input and output variables. The findings of Ahmadzadeh and Lundberg [75] indicate a remarkable level of correlation between these input and output variables. Additionally, Saldaña et al. [80] apply ML algorithms (MR, DT, ANN) to model and improve the production and power of the SAG mill, finding better operating conditions that increase or maintain production while decreasing or maintaining power.

However, the contrast between deterministic and stochastic models such as BN has not confirmed a significant advantage of neural networks compared to simpler methods [81]. With the rise of ML tools available as part of software packages, data-driven modeling applications are likely to become more prevalent and employ more sophisticated techniques and analysis. In this context, BNs are selected not for maximizing numerical prediction accuracy, but for their strengths in explainability, causal structure representation, and explicit propagation of uncertainty, capabilities that are essential for understanding and managing complex industrial processes such as SAG milling.

3.3. Bayesian Networks

Bayesian networks, also known as Bayesian belief networks or probabilistic belief models, are probabilistic models that represent causal relationships between variables and use Bayesian probability theory to make inferences. In a BN the nodes of the graph represent random variables and the edges represent probabilistic relationships between them. Each node is associated with a conditional probability distribution that describes

the probability that the corresponding variable takes a certain value, given the value of the variables with which it is directly related (in the graph) [82]. Bayesian inference has the ability to estimate the posterior probability of unknown variables based on known variables, providing interesting information about how the variables in the domain are related (which in some cases can be interpreted as cause–effect relationships) [22].

Bayesian networks provide a graphical representation of a set of random (and pseudo-random) variables and the relationships between them. The network structure allows the joint probability function of these variables to be specified as the product of conditional probability functions, which are generally simpler [83]. BNs are probabilistic and multivariate models that relate a set of random variables through a directed graph that explicitly indicates causal influence, which is possible thanks to their probability update engine: Bayes' Theorem [84]. In this sense, BNs are causal networks.

Then, defining the fundamentals of Bayesian networks, given a vector of random variables $X : (x_1, \dots, x_n)$, a joint probability measure $\Pr$ is defined as $\Pr : \text{dom}(X) \rightarrow [0, 1]$ where $\text{dom}(X) = \text{dom}(x_1) \times \dots \times \text{dom}(x_n)$. If the joint probability is known, it is possible to calculate any probability over the variables $x_1, \dots, x_n$. Then, propositions such as the rule of total probability (see Equation (1)) and the marginalization rule (see Equation (2)) are defined.

$$\Pr(X|Y) = \frac{\Pr(X, Y)}{\Pr(Y)} \quad (1)$$

$$\Pr(A) = \sum_{i \in I} \Pr(A, B_i), \quad B_i \text{ disjoint}, \quad \bigcup_{i \in I} B_i = \Omega \quad (2)$$

Then, the theorem shown in Equation (1) (rewritten in Equation (3)) shows a simple but powerful relationship between conditional probabilities, which will be the basis of the Bayesian network theory.

$$\Pr(C = c|X = x) = \frac{\Pr(X = x|C = c) \times \Pr(C = c)}{\Pr(X = x)} \quad (3)$$

where $\Pr(C = c|X = x)$: Later, $\Pr(X = x|C = c)$: verisimilitude, $\Pr(C = c)$: Prior, and $\Pr(X = x)$: Evidence. Then, considering the independence between the factors, the two random variables are independent if and only if the conditions of Equation (4) are met, or the condition of Equation (5) is met if the existence of evidence is considered.

$$\forall x, y \Pr(X = x, Y = y) = \Pr(X = x) \times \Pr(Y = y) \quad (4)$$

$$\Pr(X, Y|E) = \Pr(X|E) \times \Pr(Y|E) \quad (5)$$

If, on the other hand, X and Y are independent, $\Pr(X|Y) = \Pr(X)$ o $\Pr(X|Y, E) = \Pr(X|E)$. The independence between variables allows us to reduce the complexity of the joint probability function, and instead of modeling a single function, we separate them into simpler parts. Assuming that we have data of the form $(X_1, \dots, X_2, C)$, where C is the class variable, and the value of the class $(x_1, \dots, x_2)$ is sought to be predicted, a probabilistic approach will assign the most probable class (see Equation (6)).

$$\bar{c} = \arg \max_c \Pr(C = c|X_1 = x_1, \dots, X_n = x_n) \quad (6)$$

Then, if Bayes' Theorem is applied, it is updated $\bar{c}$ (see Equation (7)):

$$\bar{c} = \arg \max_c \frac{\Pr(X_1 = x_1, \dots, X_n = x_n|C = c) \times \Pr(C = c)}{\Pr(X_1 = x_1, \dots, X_n = x_n)} \quad (7)$$

However, to estimate the response, the expected value of mineral recovery through the product between the outputs and their conditional probabilities is considered, as shown in Equation (8).

$$E[Y(t)|X] = \sum_{i=1}^n y_i(t) \times P(y_i(t)|X), \forall t \quad (8)$$

where X represents the set of independent variables that influence the response variable to different degrees, and y_i corresponds to a possible recovery state at a given time t . If the existence of certain independent variables is unknown, the expected value of the probability of each value of the response variable is considered conditional on the evidence that the $n - k$ independent variables are known and the conditional distributions of the k independent variables are unknown, as shown in Equation (9).

$$E[Y(t)|X_{n-k}] = \sum_{i=1}^n y_i(t) \times P(y_i|X_{n-k} \wedge E[x_1, x_2, \dots, x_k]), \forall t \quad (9)$$

It should be noted that the expected value of the k variables whose evidence is unknown may or may not be conditioned by the other $n - 1$ independent variables, which are the n_k variables whose evidence is known and the conditional distributions of the $k - 1$ variables whose evidence is unknown.

The use of BN in this study is motivated by both methodological and operational considerations. From a methodological standpoint, discretization enhances the robustness of structure learning by reducing sensitivity to outliers, avoiding the need to impose linear or Gaussian assumptions, and producing more stable and interpretable dependency structures in highly variable industrial environments. From an operational perspective, SAG plant decision making commonly relies on threshold-based categories, making discrete representations more consistent with how process conditions are evaluated and managed in practice. Consequently, the primary objective of the proposed Bayesian network is not exact point regression, but rather causal analysis and scenario-based probabilistic reasoning to support uncertainty-aware operational decisions.

3.4. Structural Constraints: Blacklists and Whitelists

To incorporate process knowledge and avoid learning implausible causal directions, the structure learning algorithm was constrained using expert-defined whitelists (see Table 2) and blacklists (see Table 3). Blacklists are used to prevent edges entering exogenous variables (Hardness, SAG Rotational Speed, SAG Water Feeding, Granulometry > 100 mm, Granulometry < 30 mm, and Liner Age), as these represent upstream physical or geological drivers and should not be influenced by downstream operational variables. Furthermore, outgoing edges from the target nodes (SAG Production and SAG Power) were prohibited, ensuring that these variables behave as sinks in the causal graph, consistent with their role as system outcomes rather than system drivers.

Table 2. Whitelist—physically justified arcs.

Justification	Father Node	Son Node
Feeding and granulometry	Granulometry > 100 mm	P ₈₀ in the feeding
	Hardness	P ₈₀ in the feeding
	Stock Pile Level	P ₈₀ in the feeding
	Granulometry > 100 mm	Pebbles
	Granulometry < 30 mm	Pebbles

Table 2. Cont.

Justification	Father Node	Son Node
Solids and water	SAG Water Feeding	Solids in the feeding
	P ₈₀ in the feeding	Solids in the feeding
Hydraulic condition	SAG Water Feeding	SAG Pressure
	Solids in the feeding	SAG Pressure
	Pebbles	SAG Pressure
	SAG Pressure	Sump Level
	SAG Water Feeding	Sump Level
Influences on responses	SAG Rotational Speed	SAG Power
	SAG Rotational Speed	SAG Production
	Solids in the feeding	SAG Production
	P ₈₀ in the feeding	SAG Production
	SAG Pressure	SAG Power
	SAG Pressure	SAG Production
	Liner Age	SAG Power
	Liner Age	SAG Production
	Pebbles	SAG Production

Table 3. Blacklist—physically justified arcs.

Justification	Father Node	Son Node
Targets such as sinks	SAG Power	$\forall X \neq \text{SAG Power}$
	SAG Production	$\forall X \neq \text{SAG Production}$
Exogenous roots without parents (no node/variable in the model can “explain” hardness, degree of wear, coarse/fine particle size, etc.)	$\forall X \neq \text{Hardness}$	Hardness
	$\forall X \neq \text{Granulometry} > 100 \text{ mm}$	Granulometry > 100 mm
	$\forall X \neq \text{Granulometry} < 30 \text{ mm}$	Granulometry < 30 mm
	$\forall X \neq \text{Liner Age}$	Liner Age
	$\forall X \neq \text{Stock Pile Level}$	Stock Pile Level
	$\forall X \neq \text{SAG Rotational Speed}$	SAG Rotational Speed
	$\forall X \neq \text{SAG Water Feeding}$	SAG Water Feeding

Whitelists were used to ensure that known physical dependencies remained admissible during the search process (e.g., the possibility of links among granulometry variables, solids concentration, and water addition). These constraints reduce the search space, avoid spurious causal directions, and reflect operational knowledge of SAG mill behavior. Additionally, a maximum in-degree of 6 parents (5 in case of sinks) per node was imposed to prevent combinatorial explosion in CPT size and to ensure interpretability and data support for all conditional probabilities.

Additionally, the rules obtained from expert knowledge are listed below (see relationships R₁ to R₄).

1. R₁: f: {Solid sporcentaje in feeding, SAG RPM, SAG wáter in feeding, Hardness, Liner Age} → {SAG production}

2. $R_2: f: \{\text{Solids porcentaje in feeding, SAG RPM, SAG pressure, Hardness, Liner Age}\} \rightarrow \{\text{SAG power}\}$
3. $R_3: f: \{\text{Granulometry} < 30 \text{ mm, Granulometry} > 100 \text{ mm}\} \rightarrow \{P_{80}\}$
4. $R_4: f: \{\text{Hardness, Pebbles}\} \rightarrow \{\text{SAG pressure}\}$

These constraints operationalize expert rules R_1 – R_4 as hard admissibility conditions, ensuring that the Hill-Climbing search explores only structures consistent with established process causality rather than unconstrained statistical associations.

3.5. Validation Through Performance Measures

After developing the models, it is important to validate them using various techniques. The experiments will generate data that will allow the generation of a confusion matrix, which facilitates the analysis of classification errors. With this matrix, the performance values necessary to evaluate the classifier implementation can be calculated. The confusion matrix is a 2×2 matrix with four numerical values: TP, FP, TN, and FN, which represent the classified cases. TP is the sum of true positive cases, FP is the number of false positives, TN represents true negatives, and FN corresponds to false negatives [85].

The quality of the predictive models developed in this study is determined by merit measures based on data from the confusion matrix and the training results (see Equations (10)–(16)). The evaluation of the recommendation system result, compared to the planning generated by historical and expert data, is assessed by performance indicators such as accuracy, precision, recall, specificity, F_1 score, Matthew's Correlation Coefficient (MCC) and Kappa Index (κ), calculated from the information provided by the confusion matrix [86].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

$$\text{Specificity} - \text{TN Rate} = \frac{TN}{TN + FP} \quad (13)$$

$$F_1 \text{ score} = \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

$$\text{MCC} = \frac{TN \cdot TP - FP \cdot FN}{[P]^{0.5}} \quad (15)$$

$$\kappa = \frac{P_0 - P_e}{1 - P_e}; P_e = \frac{P}{N^2}; P = (TN + FN) \cdot (FP + TP) \cdot (TN + FP) \cdot (FN + TP) \quad (16)$$

P_0 represents the proportion of observed agreements (or accuracy) and P_e is proportion of agreements expected in the hypothesis of independence among observers, that is, agreements at random. Additionally, the ROC curve, a graphical tool used to evaluate the performance of a binary classifier, was plotted. It plots the True Positive Rate (sensitivity or recall) on the y-axis against the False Positive Rate (1–Specificity) on the x-axis at different decision thresholds. ROC curve shows the trade-off between correctly identifying positive cases and incorrectly labeling negative cases as positive. A model with perfect discrimination has a curve that passes through the upper-left corner, while a random model lies along the diagonal line. The Area Under the Curve (AUC) is commonly used as a summary measure of classifier performance, with values closer to 1 indicating better accuracy.

Finally, for discretized targets, all regression metrics (RMSE, MAE, MAPE, R^2) were computed using the expected value of the posterior distribution. Each categorical state

was assigned an ordinal numerical code, and the predicted value corresponded to the posterior expectation $E[\text{state}]$. This is a standard evaluation method for discrete or ordinal Bayesian network regressors and maintains the ordinal structure of the target variable. To evaluate the model's performance on the original continuous scale, a numerical estimate was reconstructed from the discretized output of the BN. In the first stage, a continuous representative value was calculated for each category. Then, for each observation, the posterior probability distribution over the bins was obtained, and a continuous expected value was calculated as a weighted combination of the representative values for each category. Finally, this expected value was compared with the continuous observed value of the target variable, and the R^2 , RMSE, and MAE statistics were calculated for these observed–predicted pairs on their original physical scale. These metrics provide complementary insight into the model's behavior relative to the underlying continuous variables.

4. Results

4.1. Explanatory Analysis

Spearman correlation (see Figure 3) indicates the range of correlation between the main sampled variables, which verifies logical assumptions about the working dynamics of the SAG milling process, such as the negative correlation between P_{80} and the percentage of granulometry less than 30 mm, or the positive correlation with the percentage of granulometry greater than 100 mm. The correlation study also indicates the existence of a strong positive correlation between water in the feed and TpH production. Furthermore, the lack of Spearman correlation does not indicate that there is no relationship between the variables, but rather that the relationship is not monotonic [87].

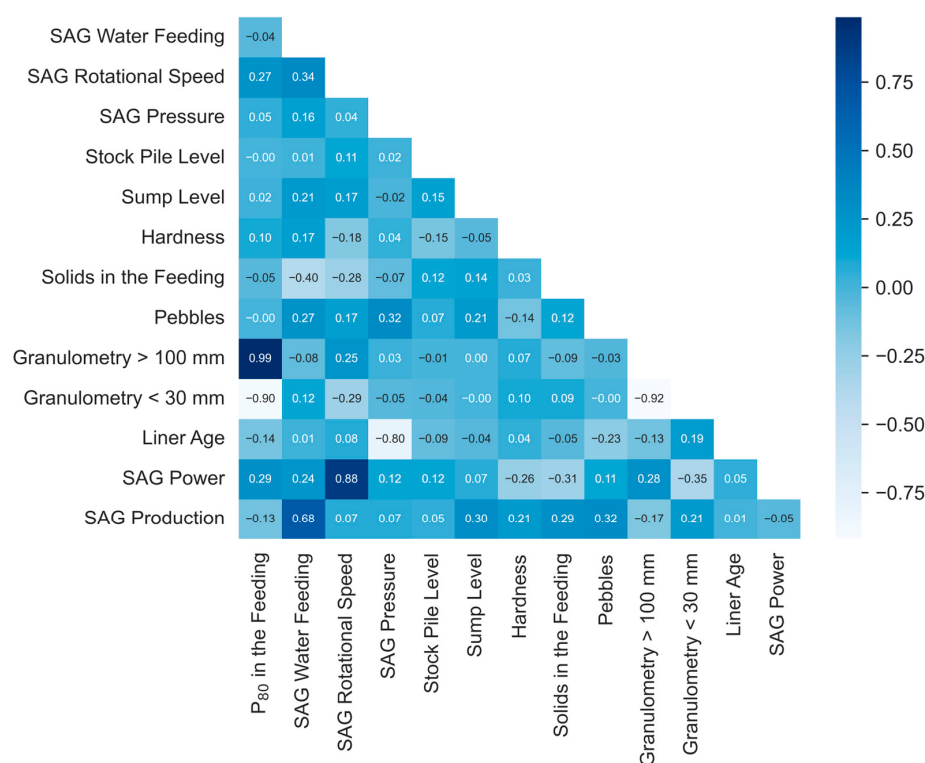


Figure 3. Spearman correlation of the SAG mill operating variables.

4.2. Discretization Strategy and Distribution Fitting Validation

4.2.1. Discretization Strategy

The analysis of the SAG milling process behavior involves a complex and multidimensional system in an environment with uncertainty and the presence of certain risk

factors involved in the problem. Modeling, enabled by probability distributions under uncertainties, provided by BN-based models is appropriate for modeling the production and power of the SAG mill, since it allows hypotheses to be deduced and relationships to be established between the explanatory factors of the response variable.

For both SAG Power and SAG Production, the Freedman–Diaconis (FD) rule suggests a very fine-grained partition (≈ 103 and ≈ 67 bins, respectively; see Table 4). Such resolutions would lead to extremely large conditional probability tables and sparsity in several parent–child combinations. Therefore, a discretization into 5 bins (very low, low, normal, high, very high) was adopted, for SAG Power and SAG Production, which provides sufficient resolution to distinguish operational regimes while ensuring robust statistical support and stable parameter estimation. From an operational perspective, SAG mills are characterized by a limited number of well-defined nonlinear operating regimes rather than a continuum of different states. Discretizing each response variable into five categories provides a parsimonious representation of these physically meaningful regimes, capturing critical transitions while preserving statistical robustness and interpretability. This level of discretization avoids unnecessary fragmentation of operational states and aligns with how SAG mill performance is monitored and controlled in practice.

Table 4. Discretization diagnostics for SAG Power and SAG Production.

Variable	n	Q1	Q3	IQR	FD Bin Width	Range	FD Suggested Bins	Chosen Bins	Obs/Bin
SAG Power	8253	19,807	22,171	2363	234	23,957	103	5	~ 1650
SAG Production	8253	3146	3808	661	65	4333	67	5	~ 1650

On the other hand, independent variables were discretized into three ordered levels (low, medium, and high) using a quantile-based binning strategy supported by the FD heuristic to ensure balanced state support and avoid sparsity in the conditional probability tables. Because these variables are treated as explanatory inputs rather than prediction targets, a coarse granularity is sufficient to represent their operational variability while preserving numerical stability in both structure learning and parameter estimation. The operational averages of all variables, together with their discretization schemes, number of bins, and corresponding value ranges, are summarized in Table 5.

Table 5. Independent operating variables of the SAG grinding process.

Variable	Unit	Bins	Mean Value	Ranges
P ₈₀ [x ₁]	mm	3	98.9065	(0, 90], (90, 105], (105, +∞)
SAG water feeding [x ₂]	m ³ /h	3	1325.4114	(0, 1200], (1200, 1400], (1400, +∞)
SAG rotational speed [x ₃]	RPM	3	8.7150	(0, 8.5], (8.5, 9], (9, +∞)
SAG pressure [x ₄]	kPa	3	7679.3956	(0, 7600], (7600, 7800], (7800, +∞)
Stockpile level [x ₅]	m	3	26.1888	(0, 20], (20, 30], (30, +∞)
Sump Level [x ₆]	m	3	89.2064	(0, 85], (85, 95], (95, +∞)
Hardness [x ₇]	-	3	35.3518	(0, 32.5], (32.5, 37.5], (37.5, +∞)
Solids in the feeding [x ₈]	%	3	71.7799	(0, 67.5], (67.5, 72.5], (72.5, +∞)
Pebble [x ₉]	TpH	3	410.7058	(0, 300], (300, 500], (500, +∞)
Granulometry > 100 mm [x ₁₀]	%	3	19.2524	(0, 15], (15, 22.5], (22.5, +∞)
Granulometry < 30 mm [x ₁₁]	%	3	39.1673	(0, 35], (35, 40], (40, +∞)
Liner Age [x ₁₂]	Month	3	3.8823	(0, 3.0], (3.0, 5.0], (5.0, +∞)
SAG Power [y ₁]	MW	5	20,820.9266	(0, 19,000], (19,000, 20,000], (20,000, +∞)
SAG Production [y ₂]	TpH	5	3443.5385	(0, 3000], (3000, 3400], (3400, 3600], (3600, 3900], (3900, +∞)

From an operational standpoint, the discretization thresholds adopted in this study are not arbitrary statistical cutoffs but are intended to represent physically meaningful operating regimes of the SAG mill. In particular, the rotational speed categories implicitly relate to sub-critical, near-critical, and supra-critical regimes, which are known to govern charge motion, impact dynamics, and energy dissipation efficiency. Likewise, the discretization of solids percentage in the feeding reflects distinct pulp density regimes, ranging from dilute transport-dominated conditions to high-density regimes associated with increased viscosity, slurry pooling, and reduced discharge efficiency. These regimes are well established in SAG mill operation and control practice and are here represented at an abstracted ordinal level (low–medium–high), rather than through explicit mechanistic thresholds, consistent with the objective of capturing dominant operational behaviors while preserving model parsimony and robustness.

In this case, the discretization procedure works primarily as a practical modeling layer that enables consistent and stable structure learning while preserving the essential transitions between operating states. This representation facilitates the integration of heterogeneous process information into the Bayesian framework and supports the scenario-based inference tasks that guide the analysis of SAG mill behavior under uncertainty.

4.2.2. Distribution Fitting Validation

The analysis of SAG milling process behavior involves a complex, multidimensional system in an environment with constant uncertainty and the uncertainty of certain risk factors involved in the problem. The best-fit distribution for the model variables is shown in Tables 6 and 7, while the graphical representation of these distributions is presented in Figures 4 and 5, for the independent and dependent variables, respectively. For continuous process variables, we considered a compact set of distributions commonly used in industrial applications: Normal, t—Student, Generalized Hyperbolic (GH), Gamma, and Weibull. Variables exhibiting substantial skewness or heavy tails were additionally tested against Skew-Normal and, when required, Johnson SU distributions. The best-fitting distribution for each variable was selected based on a combination of goodness-of-fit statistics (KS, AD), Log-likelihood, and information criteria (AIC/BIC).

Table 6. Discretization of operational variables of the SAG milling process (KS: Kolmogorov–Smirnov; AD: Anderson–Darling; Log-likelihood; AIC/BIC) [*]: Best-fitting distribution according to the corresponding goodness-of-fit statistic].

Var.	Ind.	Distribution						
		Normal	t—Student	GH	Gamma	Weibull	Skew Normal	Johnson _{SU}
x_1	KS	0.016585	0.020223	0.01795	0.024821	0.01491 *	0.017279	0.017256
	AD	3.760762	7.143328	3.293659	7.531065	3.189279	4.280355	3.062996 *
	Log-lik.	−32,431.25	−32,412.5	−32,390.21 *	−32,474.34	−32,413.92	−32,429.29	−32,391.15
	AIC/BIC	0.999785	0.999677	0.999461 *	0.999678	0.999677	0.999677	0.999569
x_2	KS	0.021705	0.021705	0.009818 *	0.019532	0.014746	0.023457	0.011884
	AD	7.167178	7.167166	0.963199 *	4.237795	1.854724	7.434172	1.400765
	Log-lik.	−54,720.04	−54,720.04	−54,677.38 *	−54,740.33	−54,700.75	−54,716.03	−54,681.69
	AIC/BIC	0.999872	0.999809	0.999681 *	0.999809	0.999809	0.999809	0.999745
x_3	KS	0.164566	0.135409	0.043681	0.149112	0.158222	0.103099	0.040372 *
	AD	375.645015	167.005365	37.705976	160.843722	188.164438	255.011176	29.311883 *
	Log-lik.	−7341.14	−5569.48	−3803.62 *	−5942.19	−5975.13	−4504.76	−4052.9
	AIC/BIC	0.99905	0.998124	0.995437 *	0.998241	0.998251	0.997682	0.996569
x_4	KS	0.045868	0.045868	0.043033 *	0.068434	0.067292	0.050537	0.04338
	AD	29.650513	29.650531	22.09348 *	51.982616	40.119984	35.020186	25.489655
	Log-lik.	−53,329.24	−53,329.24	−53,214.38	−53,376.84	−53,269.39	−53,309.54	−53,191.06 *
	AIC/BIC	0.999869	0.999804	0.999672 *	0.999804	0.999803	0.999804	0.999738
x_5	KS	0.100596	0.099831	0.020715 *	0.102747	0.099066	0.027639	0.030904
	AD	188.433358	184.918713	8.920261	142.321182	168.687387	8.153442 *	24.071989
	Log-lik.	−27,480.07	−27,477.57	−26,109.03	−27,544.08	−27,485.01	−26,031.55 *	−26,305.14
	AIC/BIC	0.999746	0.999619	0.999332 *	0.99962	0.999619	0.999598	0.999469

Table 6. Cont.

Var.	Ind.	Distribution						
		Normal	t—Student	GH	Gamma	Weibull	Skew Normal	Johnson _{SU}
x_6	KS	0.108423	0.079086	0.422935	0.080595	0.087883	0.066732	0.012207 *
	AD	199.995576	125.542184	22,172.56729	106.380384	122.78821	60.439987	2.247726 *
	Log-lik.	−26,806.61	−26,394.24	−25,646.6	−26,319.95	−26,332.28	−25,972.24	−25,644.26 *
	AIC/BIC	0.99974	0.999603	0.99932 *	0.999602	0.999602	0.999597	0.999456
x_7	KS	0.056386	0.05826	0.02149 *	0.041761	0.043112	0.030718	0.022254
	AD	44.392028	49.376112	3.612349 *	36.916985	36.666484	13.244529	4.208556
	Log-lik.	−20,122.3	−20,022.28	−19,809.88	−19,978.26	−19,971.17	−19,959.54	−19,805.01 *
	AIC/BIC	0.999653	0.999477	0.99912 *	0.999476	0.999476	0.999476	0.999295
x_8	KS	0.098829	0.086667	0.454278	0.098412	0.101752	0.059273	0.050136 *
	AD	136.126504	71.611043	31,180.31987	78.475707	82.406033	41.672367	37.54826 *
	Log-lik.	−20,928.6	−20,202.32	−19,426.85 *	−20,336.28	−20,395.75	−19,528.32	−19,507.88
	AIC/BIC	0.999666	0.999482	0.999102 *	0.999485	0.999487	0.999464	0.999285
x_9	KS	0.03352	0.03352	0.97102	0.033227	0.038529	0.019844	0.019243 *
	AD	14.789109	14.789116	109,648.6016	22.919554	17.10692	3.703869	3.651643 *
	Log-lik.	−52,062.28	−52,062.28	−51,990.36 *	−52,271.99	−52,157.25	−51,998.13	−51,994.76
	AIC/BIC	0.999866	0.999799	0.999664 *	0.9998	0.999799	0.999799	0.999731
x_{10}	KS	0.045209	0.042058	0.009295 *	0.037231	0.033572	0.014543	0.012592
	AD	42.092397	31.086192	0.613646 *	30.321545	26.723171	1.217411	0.938973
	Log-lik.	−24,969.2	−24,888.14	−24,664.96 *	−24,938.69	−24,903.69	−24,681.5	−24,668
	AIC/BIC	0.99972	0.999579	0.999293 *	0.99958	0.99958	0.999576	0.999434
x_{11}	KS	0.024616	0.029067	0.019106	0.027416	0.018637	0.018498	0.015535 *
	AD	11.006665	16.38656	2.965513 *	11.365011	7.409345	4.576473	3.347478
	Log-lik.	−26,041.56	−25,996.79	−25,896.68	−26,026.41	−25,981.51	−25,922.14	−25,894.77 *
	AIC/BIC	0.999732	0.999597	0.999326 *	0.999598	0.999597	0.999596	0.999461
y_1	KS	0.082501	0.08838	0.018689	0.080549	0.079383	0.040279	0.017925 *
	AD	157.251036	168.658464	4.345249 *	86.735191	81.153245	19.850331	4.700906
	Log-lik.	−71,703.95	−71,449.52	−70,357.85 *	−71,236.62	−71,260.35	−70,500.56	−70,361.63
	AIC/BIC	0.999903	0.999853	0.999752 *	0.999853	0.999853	0.999851	0.999802
y_2	KS	0.072416	0.073819	0.01184 *	0.038911	0.043172	0.027981	0.0131
	AD	76.117203	76.466679	1.22653 *	39.74364	42.151594	8.050869	1.4181
	Log-lik.	−61,408.4	−61,347.59	−60,939.17 *	−61,291.7	−61,271.55	−60,973.53	−60,941.69
	AIC/BIC	0.999886	0.999829	0.999714 *	0.999829	0.999829	0.999828	0.999771

Table 7. Discretization of operational variables of the SAG milling process.

Variable	Distribution	Parameters
P_{80} in the feeding	GH	$p = -7.2414; a = 0.4916; b = -0.1370; loc = 99.4382; scale = 49.9715$
SAG water feeding	GH	$p = 4.3530; a = 0.0022; b = 1.7978e-06; loc = 1324.8596; scale = 0.1758$
SAG rotational speed	GH	$p = 0.9555; a = 0.0648; b = -0.0642; loc = 9.3007; scale = 0.0004$
SAG pressure	GH	$p = 5.2903; a = 0.0327; b = 0.0070; loc = 7546.7117; scale = 1.8242$
Stock Pile Level	GH	$p = 1.5858; a = 0.2254; b = -0.2159; loc = 36.2326; scale = 0.0616$
Sump Level	Johnson's S_U	$a = 1.2445; b = 1.3632; loc = 95.9270; scale = 4.9658$
Hardness	GH	$p = 2.6174; a = 0.0086; b = 0.0025; loc = 33.3946; scale = 0.0100$
Solids in the feeding	Johnson's S_U	$a = 8.0351; b = 1.9918; loc = 78.0884; scale = 0.1971$
Pebbles	Johnson's S_U	$a = -15.4486; b = 8.1617; loc = -882.4238; scale = 395.6402$
Granulometry > 100 mm	GH	$p = 4.7248; a = 0.1803; b = 0.0801; loc = 12.3309; scale = 0.2385$
Granulometry < 30 mm	GH	$p = -3.4409; a = 6.7476; b = 2.6253; loc = 34.0098; scale = 18.4681$
Liner Age	Uniform	$a = 0; b = 6$
SAG Power	GH	$p = 1.9348; a = 11.8750; b = -11.6973; loc = 24,131.1; scale = 235.1785$
SAG Production	GH	$p = -0.0555; a = 4.8193; b = -2.9690; loc = 4111.8438; scale = 769.6986$

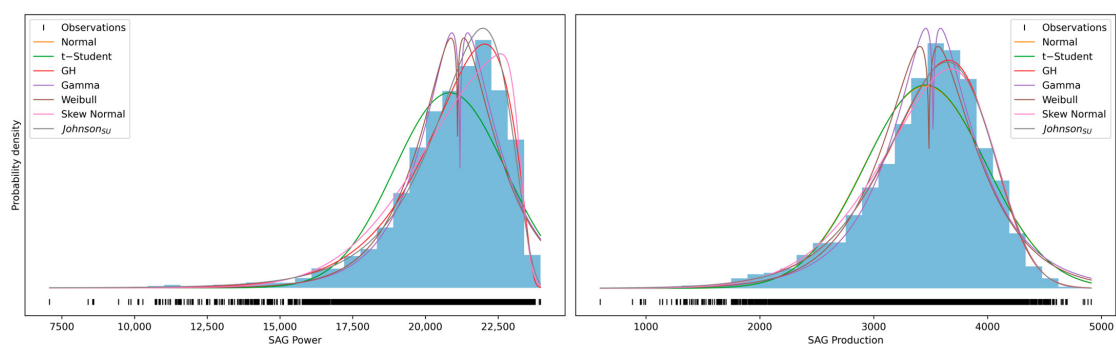


Figure 4. Distribution adjustments of dependent variables.

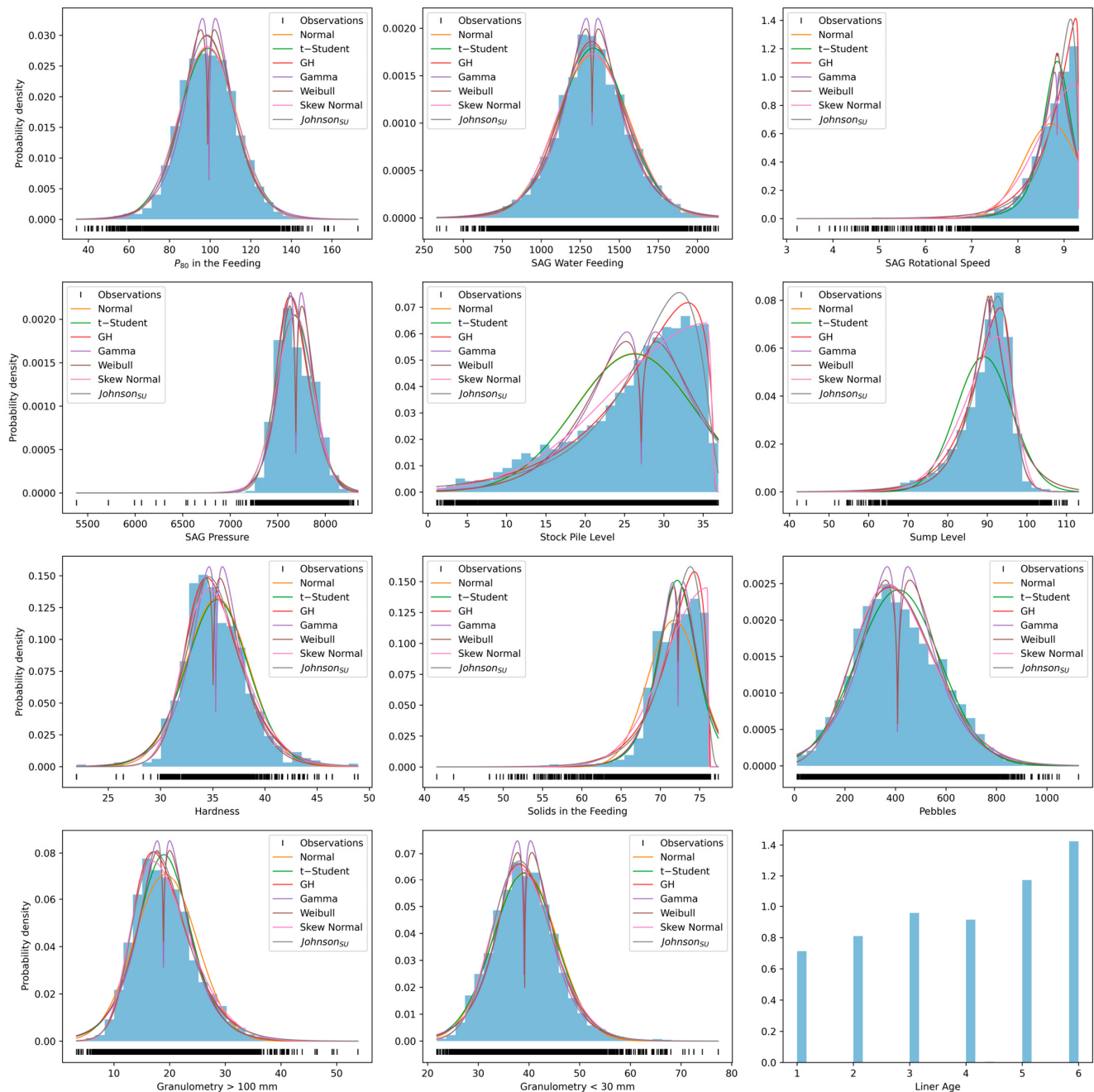


Figure 5. Distribution adjustments of the independent variables.

4.3. Bayesian Network Modeling

Building analytical models using BN inherently requires the incorporation of prior expert knowledge, particularly in complex, multidimensional industrial processes such as SAG milling, where purely data-driven associations may lead to physically implausible structures. In this study, such a priori knowledge is systematically introduced through expert-defined whitelists, blacklists, and structural rules described in Section 3.4, which constrain the space of admissible network structures and guide the learning process toward physically meaningful configurations.

In parallel, the discretization strategy described in Section 4.2.1 operates as a practical modeling layer that translates continuous process variables into ordinal operating states while preserving the essential transitions between dominant operational regimes. These discrete states are designed to represent abstracted but physically grounded behaviors,

such as sub-critical to supra-critical rotational speed regimes and low to high pulp density conditions, rather than explicit mechanistic thresholds, ensuring model parsimony and robustness. Within this constrained and physically informed framework, the structure learning process is deliberately designed to reflect known SAG mill mechanisms rather than purely statistical correlations. The enforced causal roles ensure that exogenous geological and design variables (e.g., hardness, granulometry, liner age) act as causal roots, while operational outcomes such as SAG Production and SAG Power behave as sink nodes. As a result, the learned directed acyclic graph encodes physically plausible pathways, such as the role of water and solids in governing hydraulic transport, and the influence of rotational speed and pressure on energy draw, thereby preserving causal interpretability while still allowing data-driven refinement within physically admissible boundaries.

The Discrete Bayesian Network model fit is represented as a Directed Acyclic Graph (DAG) (see Figure 6), where each node is an operational variable of the SAG circuit (discretized in ranks) and each directed edge $u \rightarrow v$ encodes the conditional dependence of v given its parents. The model factorizes the joint distribution as $P(X) = \prod_i P(X_i | Pa(X_i))$ and takes as target variables the SAG mill power and the production (ton/h), treating them as sinks (no outputs) to avoid non-causal feedbacks and facilitate their operational interpretation [88]. The DAG structure was learned with a score-based approach using Hill-Climbing with local operators of adding, deleting and reversing edges, always imposing the acyclicity constraint. A decomposable score (typically BDeu, optionally BIC/K2) was used as the objective function, which allows for efficient evaluation of local changes. Furthermore, the search space was restricted with expert knowledge (white/black lists and indegree limits) and with the decision to retain the control/exogenous variables as roots (hardness, particle size distribution, rotational speed, water in the feed, and lining age). This combination reduces overfitting, accelerates convergence, and aligns the graph with the process physics [89–93].

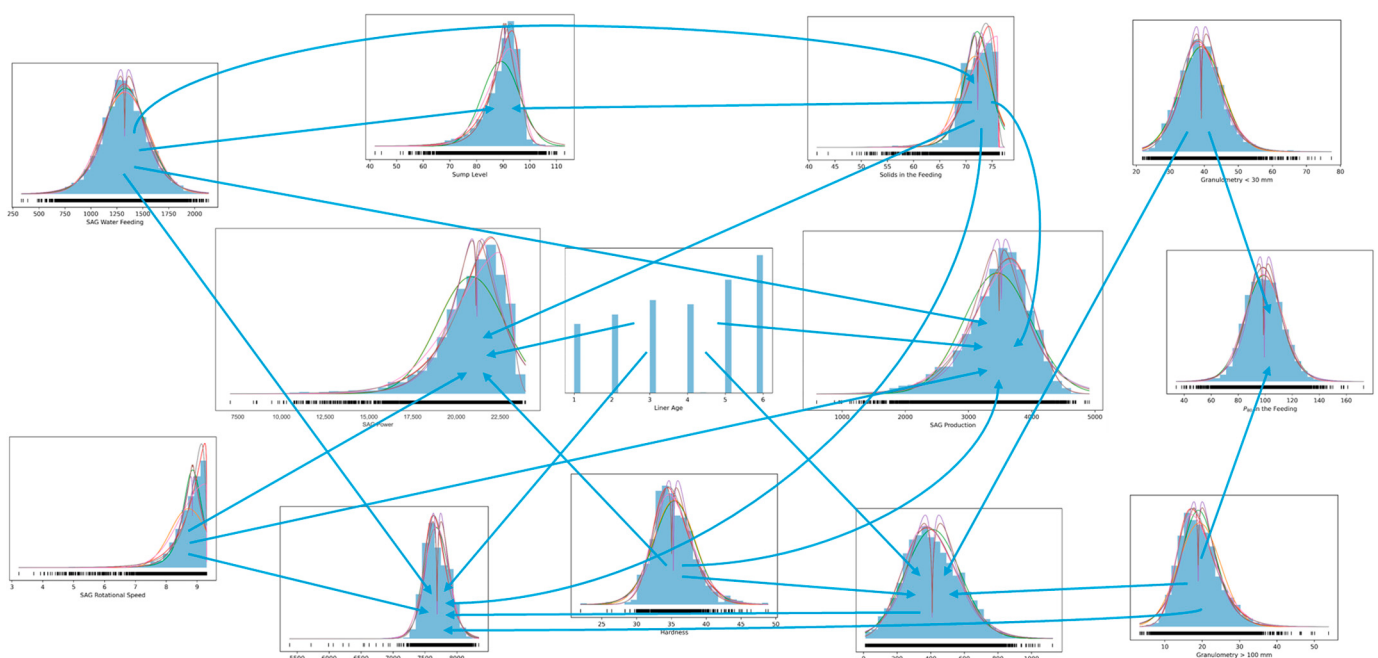


Figure 6. Bayesian network for SAG milling process.

The parameters (CPDs) were estimated with a Bayesian (Dirichlet) estimator instead of MLE, using an equivalent sample size (ESS) to smooth sparsely observed cells in the probability tables. Since real-world operations involve tails, quantile-cut discretization, and

parent–child combinations with unequal support, the Dirichlet prior acts as a regularizer, improves generalization in testing, and stabilizes inference in low-frequency scenarios (e.g., extreme bins of %solids or granulometry) [90,94]. This algorithm was chosen for three main reasons [90,91,93]:

1. Robustness and efficiency: Greedy search with decomposable scores is computationally feasible for the number of variables and states involved, unlike exhaustive (NP-hard) searches or continuous approaches that are not directly applicable to discrete variables.
2. Support for process knowledge: The score-based framework naturally integrates physical constraints (e.g., objectives such as sinks, indegree bounds, required/allowed edges), which are more cumbersome in purely constraint-based methods.
3. Out-of-sample performance: The BDeu + Dirichlet tandem penalizes overly complex models and mitigates the overfitting typical of large CPTs, resulting in better temporal validation/holdout metrics.

Finally, a discrete Bayesian network was used because the circuit signals exhibit non-normality (see Figures 4 and 5), structural zeros (e.g., pebbles), regime shifts, and non-linearities (U-shaped effects of %solids or RPM). Discretization (quantiles, specific cutoffs, and missingness handling) allows these relationships to be captured in an explainable manner and facilitates exact inference via Variable Elimination to obtain $P(\text{SAG power}|\text{evidence})$ and $P(\text{SAG production}|\text{evidence})$. Overall, the pipeline “discretization → structure learning with Hill-Climbing + score → Bayesian estimation of CPDs → exact inference” delivers an interpretable DAG, consistent with the process physics, and with a good compromise between accuracy and traceability for potential use as decision support in the plant [95,96].

Table 8 presents the conditional dependencies of each variable, including rules R₁–R₄. This alignment between the learned dependencies and the expert-imposed directional constraints confirms that the rules R₁–R₄ (along with whitelist and blacklist) shaped the admissible causal skeleton, while the data-driven search refined the specific conditional relationships within that expert-defined space.

Table 8. Summary of parent and child nodes for each node in the Bayesian network.

Node	Indegree—Parents	Outdegree—Children
Hardness	0—None	3—Pebbles, Production, Power
P ₈₀ in the Feeding	2—Granulometry < 30 mm, Granulometry > 100 mm	0—None
Solids in the Feeding	1—Water Feeding	4—Production, Power, Pressure, Sump Level
Pebbles	4—Hardness, Granulometry < 30 mm, Granulometry > 100 mm, Liner Age	1—Pressure
Granulometry < 30 mm	0—None	2—P ₈₀ in the Feeding, Pebbles
Granulometry > 100 mm	0—None	3—P ₈₀ in the Feeding, Pebbles, Pressure
SAG Production	5—Water Feeding, Solids in the Feeding, Hardness, Rotational Speed, Liner Age	0—None
SAG Power	4—Rotational Speed, Pressure, Hardness, Liner Age, Solids in the Feeding	0—None
SAG Pressure	6—Pebbles, Liner Age, Solids in the Feeding, Water Feeding, Rotational Speed, Granulometry > 100 mm	0—None
SAG Rotational Speed	0—None	3—Production, Power, Pressure
SAG Water Feeding	0—None	4—Solids in the Feeding, Production, Pressure, Sump Level
Sump Level	2—Solids in the Feeding, Water Feeding	0—None
Liner Age	0—None	4—Pebbles, Production, Power, Pressure

The network fit indicates that the stockpile inventory level does not show an effect of this exogenous variable on particle size distribution or any other process variable. This does not mean that the link between inventory level and particle size distribution (the most obvious relationship) does not exist, but rather that, with the current data and preprocessing, the signal is weak. This behavior could be explained due to low-level variability, noise, or time lags; the stockpile impacts what arrives at the mill with a delay of minutes or hours due to mixing on belts/silos, so that when comparing t with t , the algorithm sees no relationship. Furthermore, due to the multiple feeder configuration of the SAG mill under study, the process of varying the feeders and their feed rates cushions this effect, maintaining a relatively homogeneous PSD. Additionally, the literature indicates that homogenizing stacking/reclaiming practices, high live capacity, and upstream blending tend to stabilize the PSD regardless of the instantaneous level [97–99].

From the analysis of dependencies of the BN fitted, the following observations are obtained:

- Hardness → Pebbles: Greater mineral hardness increases the likelihood of accumulating particles in the hard-to-break range (≈ 25 – 55 mm), which the circuit itself recognizes as pebbles and recirculates to the crusher; therefore, it is reasonable that hardness increases the observed pebble flow rate. Dynamic and design studies of SAG circuits emphasize precisely this breaking inefficiency in this range and its management with pebble crushing [100].
- Solids in the Feeding → Sump Level: Solids percentage in the feeding governs rheology and downstream hydraulics. Higher densities imply that the sump level tends to rise (more load on pumps and cyclones), while with dilution, it falls. Control manuals and circuit models show that water/solids in the sump are a central lever for stabilizing the level and cyclone cut-off [101].
- Pebbles → SAG Pressure: Higher pebble levels increase the critical size fraction within the SAG mill, partially obstruct the passage through the grates/pulp lifters, and increase pulp residence time. This “clogging” favors the formation of slurry pooling, reduces evacuation capacity, and consequently increases the internal pressure of the SAG circuit. The literature shows that pooling is closely linked to the discharge capacity and pulp transport rate. When evacuation is limited by the accumulation of critical sizes, pressure rises and throughput drops, a phenomenon documented in industrial campaigns and mechanistic discharge analyses (grate–pulp lifter). Furthermore, recent studies on pebble recycling dynamics show that higher pebble loads increase the circulating load and aggravate pooling/clogging conditions, with a direct impact on hydraulic variables such as pressure. These mechanisms explain why pebble flow is a driver of the SAG Pressure observed at the plant [100,102,103].
- SAG Water Feeding → {Solids in the Feeding, SAG Production, SAG Pressure, Sump Level}: Water is the key manipulated variable for determining the solids percentage and, by extension, bed rheology; its adjustment impacts the sump level and pressure (transport/pooling). By improving transport and cyclone shear, it also impacts production (SAG Production) and the pebble fraction (relieving critical size retentions). Experimental and simulation evidence shows the influence of the solids percentage/water ratio on size distribution and the hydraulic performance of the circuit [104,105].
- Liner age → {SAG Production, SAG Power, SAG Pressure}: Liner age alters the lifter profile and load kinematics, simultaneously affecting throughput, power, and hydraulic conditions (e.g., alleviation or worsening of pooling). Therefore, liner age explains trends in TpH, MW, and pressure. Field and modeling work confirms that

performance often improves as the liner settles and then declines at the end of the cycle, and that liner changes modify both power and slurry evacuation [106].

4.4. Validation and Verification of the Bayesian Network

The effectiveness of the BN in estimating SAG Power and Production is evaluated, considering prior knowledge of the parameters used in the model. Prediction capability relies on knowledge of the variables of interest; otherwise, the expected value of the response variable is calculated, conditioned on the underlying distributions of the independent variables. Before estimating SAG mill output using the BN, the independent variables were considered as evidence, and the results obtained were analyzed and compared with the operational data. Various metrics (precision, accuracy, F_1 score, etc.) were evaluated (see Table 9), indicating that the model effectively represents the operational dynamics of the SAG mill.

Table 9. Bayesian network fit quality indicators for training and testing for discretized responses.

Indicator	SAG Production		SAG Power	
	Train	Test	Train	Test
Accuracy	0.844036	0.833732	0.850076	0.848313
Precision	0.609185	0.600320	0.622398	0.605359
Recall	0.578821	0.565102	0.554137	0.544076
Specificity	0.957327	0.955063	0.952465	0.952868
F_1 Score	0.571129	0.555936	0.551891	0.548892
MCC	0.668403	0.650520	0.680502	0.679226
Kappa Index	0.665144	0.646731	0.677890	0.677320
R^2	0.862866	0.85824	0.906283	0.90063
RMSE	0.465638	0.481844	0.432270	0.432403
MAE	0.206369	0.216380	0.207529	0.209084
AIC/BIC	1.680453	2.125833	1.326733	1.737242

The BN fitted shows good ordering and generalization capabilities but moderate classification power per class. For SAG Production, Accuracy decreases from 0.84 to 0.83, and for SAG Power, it decreases slightly from 0.850 to 0.848; there is a reasonable drop from train to test, indicating controlled overfitting. However, the Precision/Recall macro is low in test ($\approx 0.60/0.57$ and $0.61/0.54$, for both, SAG Production and SAG Power, respectively), indicating confusion between intermediate bins; this is consistent with what is seen in the ROCs curves, where the extremes are very well separated, and the valley is in the middle zone. The high Specificity (≈ 0.95) suggests that, on a one-versus-rest average, the model is conservative (predicting negatives well in each class) but sacrifices true positives in some classes (recall).

MCC and Kappa indices remain moderate and stable ($0.67 \rightarrow 0.65$ in SAG Production; it remains in ≈ 0.68 in SAG Power), confirming useful but not perfect performance in multiclass. In the ordinal mapping to codes, the high R^2 in test (0.86 for SAG Production; 0.88 for SAG Power) and the contained RMSE/MAE (RMSE test ≈ 0.48 and 0.45 for SAG Production and SAG Power, respectively; MAE test ≈ 0.22 for both responses) indicate that the average error in the range is less than 0.5 bins; even when it misses the exact class, it usually misses by a neighboring bin. AIC/BIC criteria are more negative in training than in testing, as expected, that is to say, the penalty increases out-of-sample, but not dramatically. Operationally, the model is very reliable in detecting extreme states (low/high), and reasonable in the center.

ROC curves for SAG Production show a highly discriminant model with a moderate and reasonable train–test gap (see Figure 7). In training, the average micro-AUC ≈ 0.984

and macro-AUC ≈ 0.957 indicate excellent separability; in testing, they drop to micro-AUC ≈ 0.981 and macro-AUC ≈ 0.950 , confirming good generalization with some overfitting (expected in a discrete BN with several variables and parents). By class, the pattern is consistent with the physics of the process: the output extremes (very low and very high) are the easiest to discriminate (AUC_{Test} ≈ 0.996 and 0.965), while the intermediate bin concentrates the loss (AUC_{Test} ≈ 0.916). This suggests overlapping operating conditions near the capacity plateau (similar combinations of %solids, RPM, pressure, and liner condition) and possibly control actions that flatten the signal in the mid-range. Still, the micro-average curve in the test (>0.98) shows that, overall, the network prioritizes the correct class well over the remaining ones. Then, from a practical perspective, the model is reliable for alerts/diagnoses when the plant drops to very low production or enters high operation; merging adjacent bins in the middle will increase the effective AUC and operational recall, and finally, if the goal is to maintain the 10-bin response, it is advisable to move discretization cutoffs in the center toward operational thresholds.

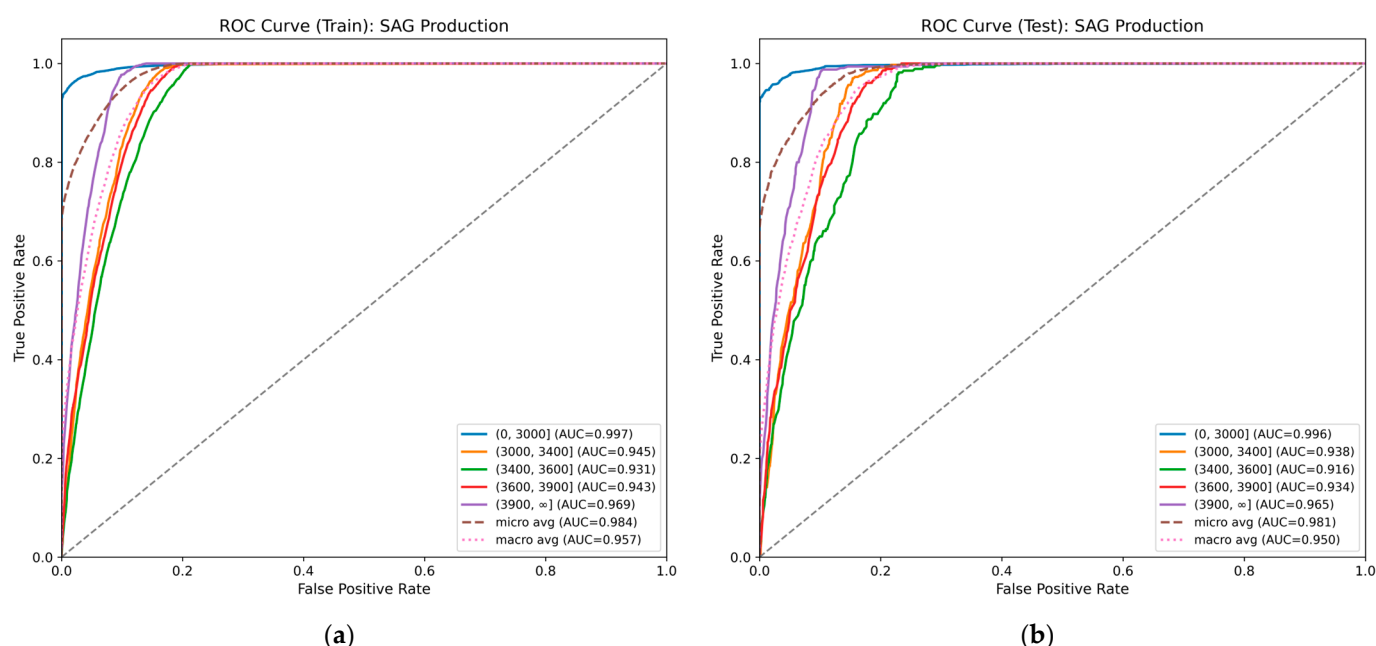


Figure 7. ROC curves for train (a) and test (b) of SAG Production [The dashed diagonal line represents the performance of a random classifier (AUC = 0.5), included as a reference baseline].

ROC curves for SAG Power (see Figure 8) show a highly discriminative model with a small generalization gap. In train, micro-AUC ≈ 0.985 and macro-AUC ≈ 0.958 are observed; in test, they drop slightly to micro-AUC ≈ 0.984 and macro-AUC ≈ 0.955 . This, as in the case of SAG Production, indicates that the BN separates power levels well and that overfitting is moderate and controlled for a discrete BN with multiple parents. By class, the power extremes are the easiest to distinguish; the lowest and highest bins reach AUC ≈ 0.996 – 0.973 in test, respectively. The middle bins show a performance valley (AUC ≈ 0.934), consistent with operational overlap near the control zone; similar combinations of SAG rotational speed, solids percentage in feeding, pressure/transport, hardness, and liner age produce similar SAG Power outputs, and control actions tend to flatten differences between adjacent classes. However, all middle bins exceed 0.91, maintaining useful ranking capability. Operationally, the model is reliable for alerts and diagnostics when power is shifted to low or high speeds, and reasonable in the midrange. If the final model uses a low/target/high decision, grouping central bins would increase the effective AUC and Recall.

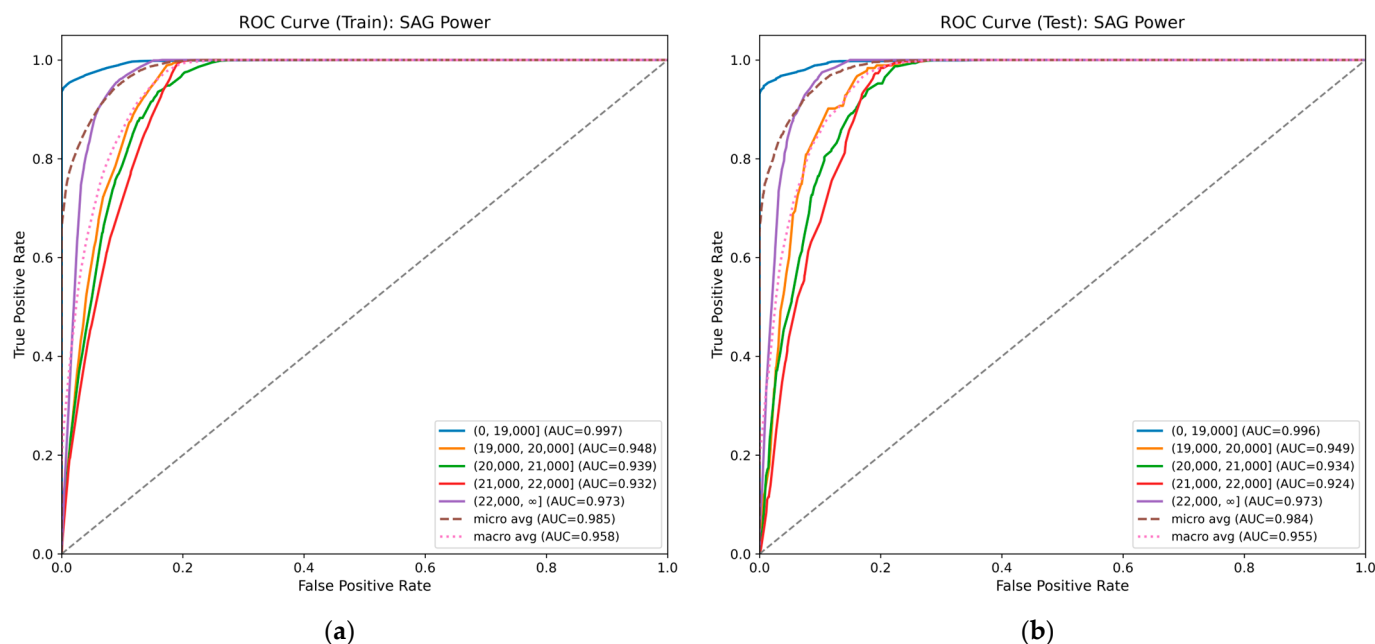


Figure 8. ROC curves for train (a) and test (b) of SAG Power [The dashed diagonal line represents the performance of a random classifier (AUC = 0.5), included as a reference baseline].

The goodness-of-fit indicators obtained for the reconstructed continuous-scale responses (see Table 10) fall within the expected range when compared with conventional regression models applied to the same database [80]. For SAG production, the BN achieves R^2 values between 0.60 and 0.63 approximately and RMSE values around 350–360 t/h, which are better than those obtained by multiple linear regression and slightly below those of ensemble-based methods such as Random Forest, XGBoost, and GBM [80]. This behavior is consistent with the intrinsic loss of granularity introduced by discretization and reflects the trade-off between interpretability and numerical precision. For SAG Power, the R^2 values around 0.60 similarly represent a reasonable performance given the higher inherent variability of this variable. Overall, the results demonstrate that although the BN is not designed as a high-precision regression engine, its continuous-scale reconstructions show stable and coherent behavior relative to established regression benchmarks.

Table 10. BN goodness of fit indicators for training and testing in continuous scale.

Indicator	SAG Production		SAG Power	
	Train	Test	Train	Test
R^2	0.629812	0.601573	0.624801	0.596750
RMSE	434.082808	449.314834	1951.403879	1999.219260
MAE	318.737717	443.473305	1443.219690	1504.043798

4.5. Structural Robustness Evaluation

The ESS sensitivity analysis performed for $ESS \in \{1, 5, 10, 20, 50\}$ using the BDeu score (Table 11) indicates a high degree of structural robustness in the learned BN. When compared with the reference model ($ESS = 10$), the Structural Hamming Distance (SHD) remains equal to zero for $ESS = 1, 5$, and 10, indicating identical network structures across this range. Minor structural differences emerge only for larger ESS values, with SHD increasing to 1 for $ESS = 20$ and to 2 for $ESS = 50$, corresponding to the addition of one or two arcs. Consistently, the number of edges remains stable at 25 for $ESS \leq 10$ and increases slightly to 26 and 27 for $ESS = 20$ and $ESS = 50$, respectively, reflecting

a modest growth in structural complexity as the equivalent sample size prior becomes more influential. In parallel, the BDeu score improves monotonically (i.e., becomes less negative) with increasing ESS, as expected when stronger priors favor denser structures and improved data fit. These results demonstrate that the inferred BN topology is largely invariant to reasonable variations in ESS, and that the learned dependencies are not driven by a particular prior specification. Considering the trade-off between structural stability, goodness-of-fit, and model parsimony, ESS = 10 was selected for the final model, as it yields an optimal balance between robustness and complexity.

Table 11. ESS sensitivity analysis.

ESS	SHD Versus Reference	Edges	BDEU Score
1	0	25	−139,313.5338
5	0	25	−137,618.8983
10	0	25	−136,929.0632
20	1	26	−136,277.3399
50	2	27	−135,560.2599

A structural bootstrap analysis with 200 resampled datasets was performed to evaluate the stability of the learned DAG. The results show a remarkably robust structure; more than 15 arcs appeared in 100% of the bootstrap replications (frequency = 1.0), including the key relationships SAG Rotational Speed → SAG Power, Solids in the Feeding → SAG Production, SAG Water Feeding → Solids in the Feeding, and Liner Age → SAG Production. Across the 200 networks, the Structural Hamming Distance (SHD) relative to the reference model was extremely low (mean = 0.79, median = 0), indicating that each bootstrap model differed by fewer than one arc on average. This demonstrates that the network topology is highly consistent under resampling, and that the directional dependencies are not artifacts of a particular sample. Overall, the Bayesian network exhibits strong structural robustness. These results are consistent with the physical behavior of SAG mills, in which the relationships involving solids content, water addition, rotational speed, granulometry, and liner condition are known to be stable determinants of production and power.

Figure 9 presents the bootstrap arc-frequency heatmap obtained from 200 resampled BNs. Several arcs display a frequency of 1.0 (yellow), meaning that they occurred in 100% of the bootstrap replications. These include key operational relationships such as SAG Water Feeding → Solids in the Feeding, Granulometry > 100 mm → P₈₀ in the Feeding, SAG Rotational Speed → SAG Power/Production, and Solids in the Feeding → SAG Production. These arcs represent structurally invariant dependencies and confirm that the core topology of the network is extremely robust. A small number of arcs show intermediate frequencies, suggesting secondary or context-dependent dependencies, while the vast majority of possible arcs exhibit zero frequency, indicating that the model does not produce spurious connections under resampling.

It is worth noting that the sparse appearance of the heatmap is an expected outcome given the domain-informed structural constraints imposed during learning (exogenous roots, sinks, and blacklist restrictions). Under such priors, stable edges tend to concentrate in well-supported operational pathways, while non-permissible relations systematically appear with zero frequency. Combined with the very low SHD obtained in the bootstrap evaluation (mean = 0.79, median = 0), this pattern reflects structural stability rather than instability, consistent with recommendations in the BN learning literature.

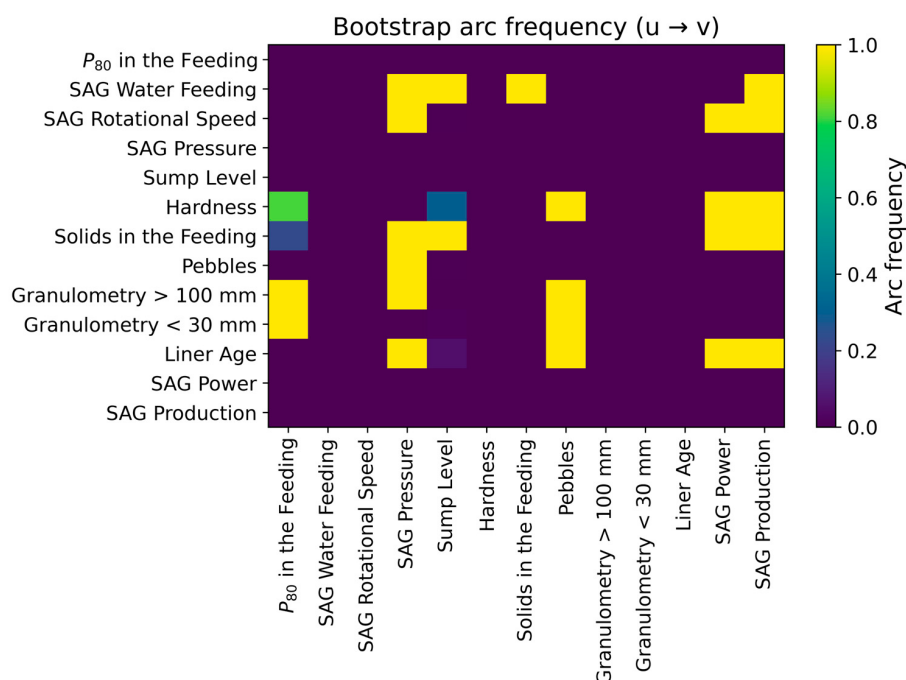


Figure 9. Bootstrap arc frequency ($u \rightarrow v$).

The comparison of Bayesian network structures learned using different scoring functions (BDeu, BIC, and K2) reveals a substantial degree of structural consistency across model selection criteria (Table 12). The structures obtained with BDeu and BIC are highly similar, with a Structural Hamming Distance (SHD) of 2 despite a small difference in the number of edges (25 versus 23), indicating that the core network topology is largely invariant to whether a Bayesian score or a penalized-likelihood criterion is employed. Comparisons involving the K2 score result in larger SHD values (SHD = 3 when compared with BDeu and SHD = 5 when compared with BIC), reflecting the fact that K2 does not penalize model complexity and therefore favors denser structures (28 edges). Nevertheless, these differences remain modest relative to the size of the search space and do not modify the principal dependency pathways identified in the network. Overall, the results demonstrate that the central causal structure is robust across different scoring functions, supporting the conclusion that the learned relationships are not artifacts of a particular score choice.

Table 12. Comparison of structure by score.

Score A	Score B	SHD	Edges A	Edges B
BDeu	BIC	2	25	23
BDeu	K2	3	25	28
BIC	K2	5	23	28

The stability of the structure across scores is consistent with the physical nature of the SAG milling process, where fundamental dependencies (speed $\rightarrow$ power, solids $\rightarrow$ production) are determined by first principles and operational constraints. Even when a more permissive score (K2) is used, these relationships persist, indicating that the BN is learning true underlying process behavior rather than score-induced byproducts.

Collectively, preventing exponential CPT growth, structural robustness study and bootstrap demonstrate that the BN exhibits strong resistance to overfitting at both structural and parametric levels. The combination of constrained in-degree, ESS stability, and bootstrap robustness establishes that the model generalizes well across different operating

regimes and is not tuned to short-term or sample-specific patterns. It is shown that the BN structure is stable across resampling and hyperparameter variation, suggesting low risk of structural overfitting. To further verify this, we performed a time-block validation to evaluate generalization across liner-wear cycles (following section).

4.6. Temporal Validation—Time-Block Split

The results of the temporal block validation are summarized in Table 13. As expected for autocorrelated SAG mill operations, a moderate degradation in predictive performance is observed when training and testing are conducted on chronologically separated data blocks. Across both SAG production and SAG power targets, classification-oriented metrics (including Accuracy, Precision, Recall, Specificity, F_1 score, MCC, and Cohen's Kappa) exhibit relative decreases in the range of approximately 5–10%. These variations are consistent with those typically reported in time-aware validations of industrial process models and reflect the increased difficulty of generalization across distinct operating periods. R^2 decreases by approximately 8–13%, while error-based metrics (RMSE and MAE) increase by roughly 10–19% in the testing blocks. Such changes are expected in SAG milling contexts, where each liner-wear cycle is associated with shifts in breakage behavior, throughput, and power consumption, leading to non-stationary input–output relationships over time. Additionally, AIC/BIC display only limited relative changes (≈ 3 –4%), indicating that neither model likelihood nor effective structural complexity deteriorates substantially when evaluated on future operational regimes.

Table 13. BN fit quality indicators for training and testing in temporal validation [$\uparrow$: increase; $\downarrow$: decrease].

Indicator	SAG Production		SAG Power	
	Train Δ	Test Δ	Train Δ	Test Δ
Accuracy	0.7963 $\sim\downarrow 5.66\%$	0.7673 $\sim\downarrow 7.97\%$	0.7845 $\sim\downarrow 7.71\%$	0.7654 $\sim\downarrow 9.77\%$
Precision	0.5751 $\sim\downarrow 5.6\%$	0.564 $\sim\downarrow 6.05\%$	0.5792 $\sim\downarrow 6.94\%$	0.5579 $\sim\downarrow 7.84\%$
Recall	0.5475 $\sim\downarrow 5.41\%$	0.5264 $\sim\downarrow 6.85\%$	0.5179 $\sim\downarrow 6.54\%$	0.4996 $\sim\downarrow 8.17\%$
Specificity	0.9029 $\sim\downarrow 5.69\%$	0.8918 $\sim\downarrow 6.62\%$	0.9025 $\sim\downarrow 5.25\%$	0.8911 $\sim\downarrow 6.48\%$
F_1 Score	0.561 $\sim\downarrow 5.5\%$	0.5446 $\sim\downarrow 6.42\%$	0.5468 $\sim\downarrow 6.73\%$	0.5271 $\sim\downarrow 8\%$
MCC	0.6298 $\sim\downarrow 5.78\%$	0.6074 $\sim\downarrow 6.63\%$	0.6359 $\sim\downarrow 6.55\%$	0.6139 $\sim\downarrow 9.62\%$
Kappa Index	0.6245 $\sim\downarrow 6.11\%$	0.5898 $\sim\downarrow 8.8\%$	0.6259 $\sim\downarrow 7.67\%$	0.6125 $\sim\downarrow 9.57\%$
R^2	0.7929 $\sim\downarrow 8.11\%$	0.7535 $\sim\downarrow 12.2\%$	0.8243 $\sim\downarrow 9.05\%$	0.7809 $\sim\downarrow 13.29\%$
RMSE	0.5143 $\sim\uparrow 10.45\%$	0.5517 $\sim\uparrow 14.5\%$	0.4959 $\sim\uparrow 14.72\%$	0.5156 $\sim\uparrow 19.24\%$
MAE	0.2284 $\sim\uparrow 10.68\%$	0.2478 $\sim\uparrow 14.52\%$	0.2355 $\sim\uparrow 13.48\%$	0.2483 $\sim\uparrow 18.76\%$
AIC/BIC	1.6293 $\sim\downarrow 3.04\%$	2.0473 $\sim\downarrow 3.69\%$	1.2805 $\sim\downarrow 3.48\%$	1.6665 $\sim\downarrow 4.07\%$

Overall, the network generalizes well across liner cycles, confirming that the learned causal relationships remain valid under time-aware evaluation rather than being byproducts of random sampling.

4.7. Interventional Scenario Analysis—What-If Conditions

To explore the operational sensitivities captured by the BN, four intervention scenarios (“what-if” conditions) were evaluated using do-calculus-based evidence manipulation (see Table 14). The baseline state was defined from the most frequent operating categories of the dataset, and each scenario modified only the variables of interest while maintaining all other conditions fixed. Expected Code represents the expected code of target bin (Power or Production), where 0 $\approx$ lowest level and 5 $\approx$ highest level, while Delta versus Baseline indicates how much the expected value rises or falls with respect to the baseline operating state.

Table 14. Results of what-if scenarios.

Scenario	Target	Expected Code	Delta Versus Baseline
Baseline	SAG Power	0.000207	0
	SAG Production	0.000014	0
Scenario 1: +RPM (alto) and baseline water/solids	SAG Power	3.756627	3.756419
	SAG Production	0.152091	0.152078
Scenario 2: +Water and -Solids	SAG Power	0.000207	0
	SAG Production	0.559289	0.559275
Scenario 3: +Hardness and +Pebbles	SAG Power	0.70227	0.702062
	SAG Production	0.040323	0.040309
Scenario 4: Near-optimal regime (high RPM, moderate solids, medium water)	SAG Power	3.679789	3.679582
	SAG Production	1.371145	1.371131

The conclusions obtained from the simulation of the scenarios are presented below:

- Scenario 1—Increasing SAG rotational speed to its highest level while keeping water addition and solid concentration at baseline levels. This intervention produced a strong rise in SAG Power ($\Delta \approx +3.8$ bins) but only a negligible change in SAG Production, reflecting that RPM primarily influences energy draw rather than throughput efficiency. The above is consistent with SAG mill physics; increasing RPM increases kinetic energy and torque demand, thus increasing SAG Power. On the other hand, increasing RPM does not guarantee greater tonnage; it can even reduce it due to cataracting or load instability.
- Scenario 2—Increasing water feeding while decreasing solids, representing a more dilute pulp regime. In this case, SAG Production increases ($\Delta \approx +0.56$), whereas SAG Power remains unchanged relative to the baseline ($\Delta \approx 0$). This behavior is consistent with improved material transport rather than increased mechanical load. The additional water does not lead to higher power consumption, which is expected since power draw is primarily governed by the mass of material being lifted and impacted inside the mill, rather than by slurry volume alone. However, it does result in a clear increase in SAG production (+0.56), as higher water addition enhances slurry mobility and a lower solids content reduces pulp viscosity, leading to improved flow conditions and faster material evacuation.
- Scenario 3—Simultaneously increasing hardness and pebbles, which moderately elevated SAG Power ($\Delta \approx +0.7$ bins) and produced a very small increase in SAG Production, a behavior aligned with the higher energy requirements imposed by tougher ore and coarse circulating load. This behavior is consistent with the operational physics of the SAG mill. In general, harder ore tends to increase the specific energy demand and, in many cases, can even reduce production due to the greater effort required for size reduction. Likewise, a higher presence of pebbles is usually associated with coarser particle sizes, which would increase the work required.
- Scenario 4—Combining high rotational speed with intermediate solids and medium water levels, representing a near-optimal operating balance. This configuration simultaneously generates a significant increase in both SAG Power ($\Delta \approx +0.56$) and SAG Production ($\Delta \approx +0.56$), reflecting a favorable balance between available energy and pulp transport efficiency. The high rotational speed increases the specific energy available for comminution, while the solids and water at intermediate levels promote a stable flow without overloading the mill. As a result, this scenario stands out as the most efficient among those evaluated, demonstrating a balanced increase in both Power and Production, and revealing operating conditions that approach optimal SAG circuit performance.

4.8. Bayesian Network Sensitivity Analysis

The univariate sensitivity analysis shown in Figure 10 indicates that SAG Power (Figure 10a) is most sensitive to SAG rotational speed. Several rotational speed states produce the largest absolute deviations from the baseline, with high-speed regimes exhibiting the strongest negative shifts (large bars on the left). This result indicates that, around the selected operating point, moving the rotational speed into these ranges substantially reduces the probability of the target SAG Power state, suggesting that the system may transition toward less efficient operating conditions at that power level. This behavior is consistent with known SAG mill dynamics, where excessive speed can lead to reduced effective grinding due to slip or incipient centrifugation effects. In contrast, low levels of SAG rotational speed and SAG pressure display the largest positive contributions (bars on the right). This suggests that, in the learned model around the baseline, pressure is acting as an indicator of operating regime rather than a direct monotonic proxy for load; therefore, the association should be interpreted as local conditional sensitivity rather than direct causality. Other variables, including P_{80} in the feeding, sump level, SAG water feeding, and particle size fractions, exhibit comparatively smaller but systematic effects.

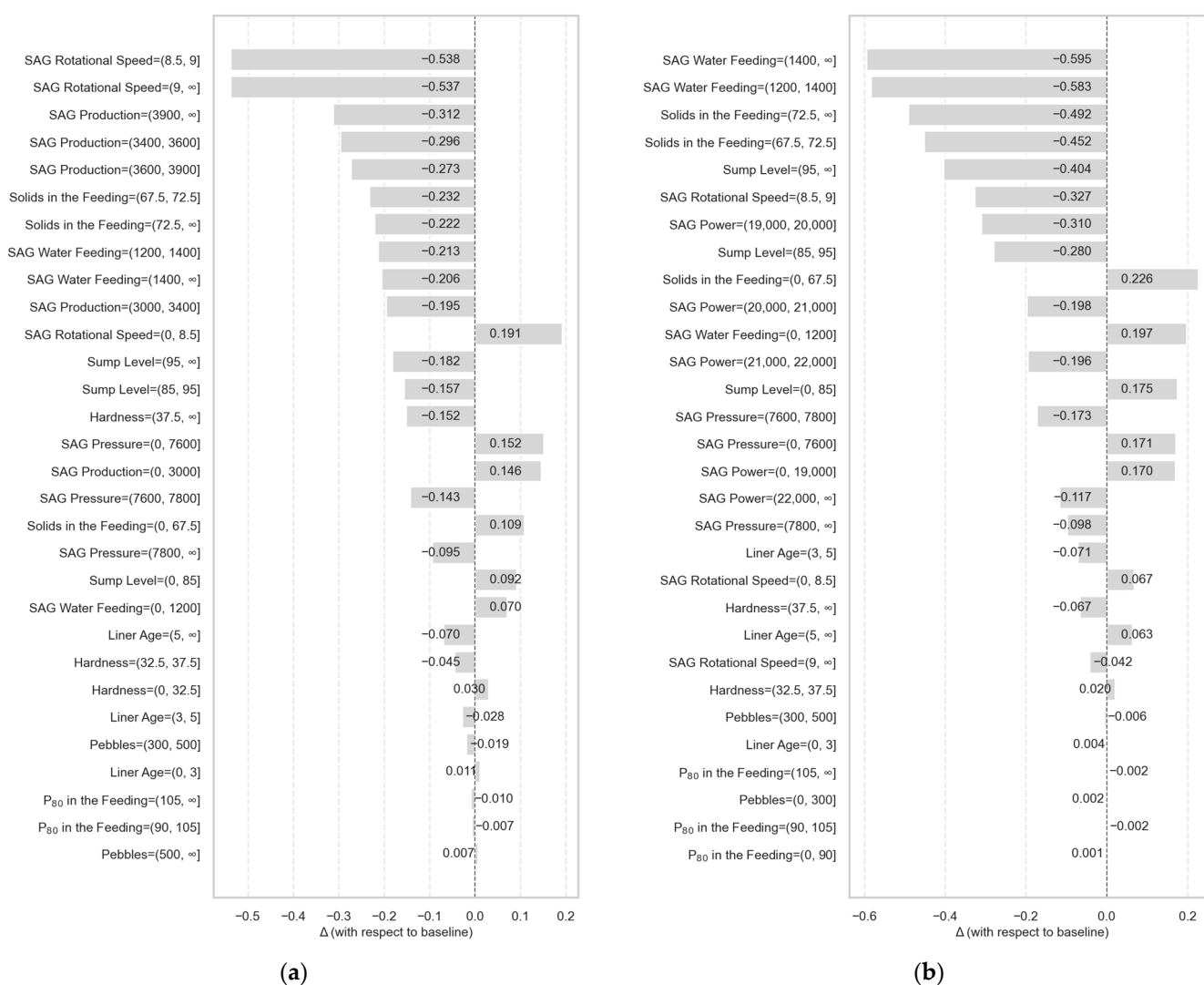


Figure 10. Univariate sensitivity analysis of the impact by state on $P(\text{SAG Power} \mid \text{baseline})$ (a) and $P(\text{SAG Production} \mid \text{baseline})$ (b).

On the other hand, SAG Water Feeding states emerge as the dominant drivers of the target SAG Production probability in the univariate SA (Figure 10b). In particular, high water feeding regimes produce the largest absolute deviations from the baseline, with upper water ranges strongly reducing the probability of the target production state, while lower water ranges exert a positive effect. This pattern is consistent with transport-dominated behavior around the selected operating point, where excessive dilution may reduce effective grinding or classification efficiency despite improving slurry mobility. These results should be interpreted as local conditional sensitivities around the selected baseline (not direct causal effects) and may not extrapolate beyond this operating region. Solids percentage in the feeding follows as a second major influence, with high-solids regimes producing marked negative shifts and lower-solids regimes increasing the probability of higher throughput. This behavior reflects the well-known trade-off between pulp density, viscosity, and transport efficiency, where excessive solids content tends to hinder material flow and limit effective production. Granulometric variables show negligible effects with physically plausible directions, as coarser feed characteristics tend to reduce throughput.

Overall, Figure 10 highlights rotational speed, pressure, and solids percentage in the feeding among the dominant local drivers of SAG Power variability around the baseline operating condition. Particle size distribution acts as negligible factors, but with lower marginal influence under the univariate perturbations considered. Additionally, SAG pressure, SAG power, rotational speed, liner age, and hardness display comparatively smaller and mixed marginal effects over SAG Production, indicating that production is primarily governed by water solids management and hydraulic transport, while mechanical and liner-related variables play a more limited role in marginal throughput variations.

5. Discussions

5.1. On the Modeled and Fitted Bayesian Network

The discrete BN fitted is capable of capturing key operational dependencies of SAG mill and generalizing with moderate out-of-sample loss, and beyond predictive behavior, the principal value of the fitted BN lies in its ability to provide an interpretable and causally coherent description of how operational factors interact under uncertainty. Because the network was learned under expert-specified white and black lists, the resulting topology reflects a hybrid integration of domain knowledge and data, thereby ensuring that the inferred causal structure adheres to physical plausibility while still capturing statistically supported dependencies. The discretized representation enhances this interpretability by structuring the process variables into meaningful operating states, enabling the network to evaluate “what-if” scenarios and quantify the likelihood of different production or power outcomes when conditions change. This capability positions the model as a decision-support tool for understanding process dynamics rather than as a purely predictive regression engine.

In terms of predictive capability, the training-to-test accuracy drop (SAG Production: $0.84 \rightarrow 0.83$; SAG Power: $0.85 \rightarrow 0.848$) remains within reasonable ranges, while MCC and Kappa remain at useful but not perfect levels (≈ 0.67 and 0.65 in SAG Production; ≈ 0.68 and 0.68 in SAG Power, for training and testing, respectively), suggesting controlled overfitting and acceptable ordering power on a multiclass problem with 5 bins in each of the responses (Production and Power). Furthermore, when mapping ordinal states to codes (testing), the high R^2 (≈ 0.90 for SAG Power and ≈ 0.86 for SAG Production) and mean errors < 1 bin (RMSE test < 0.5) indicate that the errors tend to be directed toward adjacent classes rather than distant classes, a valuable aspect for operational support where exact boundaries may be fuzzy. The AIC/BIC criteria show the expected out-of-sample penalty without evidencing dramatic degradation. These findings together outline a reliable model

for discriminating extreme states and a reasonable one in the middle zone, which makes the fitted model a powerful study and predictive tool in productive environments even when not all the input variables are certain.

ROC analysis reinforces this reading. For SAG Production, the BN shows AUCs of micro $\approx 0.98/0.98$ (train/test) and macro $\approx 0.96/0.95$, with better separability at the extreme zones and a performance valley in the intermediate bins. For SAG Power, the gap is even narrower (micro $\approx 0.99/0.98$, macro $\approx 0.96/0.95$). Operationally, this implies that the network prioritizes very low- or very high-response events quite well (which would be extremely useful for alerting/diagnosis) and that a possible aggregation of core classes could improve sensitivity without losing traceability.

From an explanatory perspective, univariate sensitivity is consistent with the physics of SAG milling. In the analyzed baseline, SAG Water Feeding dominates the impact on SAG Production (transport/hydraulics), followed by Sump Level and Solids Percentage in the Feeding; for SAG Power, rotational speed concentrates the greatest effect (with indications of inefficiency when moving outside the operating point), and SAG Pressure and Solids Percentage in the Feeding act as positive levers, while particle size and the sump show minor/modulating effects. This pattern adequately distinguishes control levers (such as water level, solids percentage, rotational speed, pressure) from more consequential or second-order variables (particle size within certain ranges), reinforcing the use of the network as an operational what-if tool. Additionally, the role of the expert rules is evident in the absence of inadmissible parent sets—for example, no operational response variable appears as a parent of upstream process conditions—demonstrating that the learned graph is consistent with the enforced causality encoded in R_1 – R_4 .

Methodologically, the combination of score-based search (Hill-Climbing) with BDeu/BIC, expert constraints, and Bayesian estimation of CPDs (Dirichlet, ESS) proved adequate to balance interpretability, robustness, and computational efficiency. This framework allows the evaluation of local changes in structure in a decomposable form, incorporating process knowledge, and smoothing sparsely observed cells, mitigating overfitting typical in large CPTs. The choice of a discrete model, given the non-normalities (shown in Figures 4 and 5) and non-linearities, also facilitated exact inference by Variable Elimination with traceability of results. A relevant observation of the DAG is the absence of a direct effect of stockpile level on other variables. This does not negate the underlying physical link but suggests a weak signal given the dataset conditions: low level variability, sensory noise/saturation, or temporal misalignment (lags) between the level and the actual feed to the mill after mixing and transport. Furthermore, the multi-feeder configuration and high live capacity stacking/reclaim practices and upstream blending tend to stabilize the PSD, dampening changes due to instantaneous level (configuration already applied in the circuit under study); this explains why the algorithm does not detect a robust dependence in this sense.

Some limitations and opportunities for improving the BN are presented below:

1. Identifying relationships with lags will require time-shifts and/or dynamic models (DBN) to capture delayed effects (e.g., stockpile $\rightarrow$ PSD $\rightarrow$ SAG Production).
2. Sensitivity showed clear levers. A next step can integrate optimization under uncertainty (e.g., recommending water/solids percentage/rotational speed setpoints) using BN as a probabilistic simulator.

Overall, the fine-tuned BN offers a compromise between accuracy, interpretability and action, suitable for diagnosis and evaluation of operational scenarios. Going a step further, the generated stochastic model could be integrated into a simulation system that allows quantifying the benefits derived from the incorporation of probabilistic models in

the estimation of the expected value of mineral recovery, or it could be used as part of a decision support system in the mining industry [107].

5.2. On Comparative Benchmarking with Alternative Machine Learning Models

The development of predictive models for SAG mill operations has increasingly leveraged advanced ML techniques alongside traditional stochastic approaches. While the BN framework presented offers distinct advantages in uncertainty quantification and causal reasoning, it is essential to contextualize its performance relative to alternative ML paradigms that have gained prominence in mineral processing applications over the past decade. This subsection provides a comparative benchmarking analysis of the BN model against commonly deployed ML approaches, with emphasis on their applicability to SAG mill power consumption and throughput prediction under operational uncertainty.

5.2.1. Random Forest and Tree-Based Ensemble Methods

RF has emerged as a robust baseline for grinding mill prediction tasks due to its inherent resistance to overfitting and capacity to handle heterogeneous industrial datasets with minimal feature preprocessing [108,109]. In grinding power prediction applications, RF has demonstrated competitive performance, with some works reporting $R^2 \approx 0.94$ for SAG mill energy consumption forecasting [15]. The ensemble averaging mechanism of RF provides natural variance reduction, making it particularly effective when dealing with noisy sensor data and incomplete feature spaces commonly encountered in industrial SAG operations [38,109].

However, RF exhibits notable limitations in the context of SAG mill modeling. Its tree-based partitioning structure inherently struggles with extrapolation beyond the training distribution, a critical concern given the dynamic nature of ore characteristics and operational regimes in mineral processing circuits [110,111]. Vera Ruiz et al. [112] reported substantially lower performance ($R^2 \approx 0.47$) for RF compared to linear regression and neural approaches in certain SAG power datasets, suggesting that RF's predictive capacity is highly dataset-dependent. Furthermore, RF lacks the explicit probabilistic framework necessary for rigorous uncertainty propagation, a capability central to the BN approach. While RF can provide prediction intervals through quantile regression forests, these intervals do not capture the joint uncertainty distributions or conditional dependencies that BNs naturally represent [113].

5.2.2. Gradient Boosting Machines and XGBoost

Gradient boosting methods, particularly XGBoost, have demonstrated superior performance in recent comparative studies of SAG mill prediction. Pural et al. [110] identified XGBoost as the most accurate method for generic mill liner wear prediction, achieving MAPE $\approx 5.27\%$ under interpolation and 6.12% under extrapolation conditions. The sequential error-correction mechanism of boosting algorithms enables effective capture of complex nonlinear residuals that persist in mechanistic models [114,115]. Integration of XGBoost with physics-based approaches has proven particularly promising. Feng et al. [115] developed a hybrid mechanistic–XGBoost model where a first-principles model captures primary physical dependencies, while XGBoost learns structured residuals, achieving superior accuracy compared to purely data-driven or purely mechanistic alternatives.

Despite these strengths, gradient boosting methods face several constraints in SAG mill applications. The sequential training process requires careful hyperparameter tuning and can be computationally intensive for large-scale industrial datasets [114]. More critically, XGBoost shares RF's limitation in extrapolation performance, with accuracy degrading substantially outside training operating regimes [110,115]. Additionally, boosting methods provide feature importance metrics but lack the explicit causal structure and bidirectional

inference capabilities inherent to BNs; that is to say, it cannot efficiently reason backward to infer likely operating conditions given observed throughput anomalies, a diagnostic capability central to the BN framework.

5.2.3. Artificial Neural Networks and Deep Learning Architectures

ANNs, and particularly recurrent architectures such as Long Short-Term Memory (LSTM) networks, have achieved state-of-the-art performance in temporal forecasting of SAG mill energy consumption. Lopez et al. [116] reported LSTM prediction errors below 4% RMSE for energy forecasting, while Avalos et al. [15] identified recurrent neural networks (RNNs) as the most accurate approach for short-term SAG energy prediction using operational variables. The capacity of LSTMs to capture temporal dependencies in sequential mill operation data provides significant advantages for real-time control applications [116,117]. Recent innovations have combined convolutional neural networks (CNNs) with attention mechanisms and physics-informed constraints. Zhang et al. [118] developed a Channel-Attention CNN-LSTM (CACN-LSTM) architecture for SAG mill power prediction, while Hermosilla et al. [119] proposed a CNN-Physics-Informed Neural Network (PINN) hybrid for overload detection, achieving F_1 scores of 94.5% with improved physical interpretability compared to purely data-driven CNNs.

However, ANNs present significant interpretability challenges that limit their adoption in decision-critical mineral processing environments. Multiple studies emphasize that despite high predictive accuracy, ANNs function as “black boxes” with limited physical interpretability [71,117,120]. This opacity complicates root-cause analysis during mill upsets and hinders operator trust, issues that motivated the development of hybrid physics-informed approaches [115,119]. Furthermore, ANNs typically require substantial training data and are prone to overfitting in small-sample regimes common in specialized ore campaigns [120,121]. Unlike BNs, which explicitly encode conditional independence structures, ANNs learn implicit representations that provide no direct insight into causal relationships between mill variables.

5.2.4. Hybrid and Ensemble Intelligent Systems

Recent research has emphasized hybrid architectures that combine multiple ML paradigms or integrate ML with mechanistic knowledge. Ghasemi et al. [68,122] developed a comprehensive framework integrating expert knowledge, CatBoost ensemble learning, and evolutionary algorithms for SAG mill throughput optimization, reporting CatBoost as the most accurate predictor among tested ensemble methods. The integration of genetic algorithms with neural networks (GA-ANN) has enabled simultaneous parameter optimization and sensitivity analysis [57,71]. Additionally, physics-informed hybrid models represent a particularly promising direction. Beyond the mechanistic-XGBoost coupling mentioned earlier [115], researchers have explored surrogate-accelerated global sensitivity analysis [37] and multi-kernel SVM approaches [123].

Nonetheless, hybrid systems introduce additional complexity in model development, validation, and deployment. The integration of mechanistic and ML components requires careful interface design and can complicate uncertainty quantification when mechanistic model assumptions are violated [37,115]. Furthermore, while hybrids improve predictive accuracy and physical consistency, they do not inherently provide the causal inference and diagnostic reasoning capabilities of BNs.

5.2.5. Comparative Synthesis: Strengths, Limitations, and Complementary Capabilities

Table 15 summarizes the comparative performance characteristics of alternative ML models relative to the BN framework for SAG mill applications.

Table 15. Comparative assessment of machine learning models for SAG mill prediction.

Model	Predictive Accuracy	Interpretability	Uncertainty Quantification	Causal Inference	Computational Efficiency	Extrapolation Capability
Random Forest	Moderate to High (R ² : 0.47–0.94) [15,112]	Moderate (feature importance)	Limited (quantile forests)	No	High	Poor [110,111]
XGBoost GBM	High (MAPE: 5.27–6.12%) [110]	Moderate (feature importance)	Limited	No	Moderate	Moderate [110,115]
ANN/LSTM	Very High (RMSE < 4%) [116]	Low (black box)	Limited	No	Low (training-intensive)	Moderate [120]
Hybrid	Very High [115,119]	High (physics-constrained)	Moderate	Partial	Moderate	Improved [115]
Bayesian Networks	Moderate to High	High (graphical structure)	High (joint distributions)	Yes	High (inference)	Moderate

Then, comparative analysis reveals the following key insights:

1. In terms of pure predictive accuracy for point forecasts, gradient boosting methods and deep neural networks consistently achieve the highest performance metrics in recent benchmarking studies [15,68,110,116]. These methods excel in capturing complex nonlinear patterns and temporal dependencies in large datasets, making them highly suitable for real-time forecasting and control applications where predictive precision is paramount [116,117].
2. Regarding interpretability and physical consistency, tree-based methods provide feature importance rankings and hybrid physics-informed models enforce physical constraints [115,119]; only BNs offer explicit graphical representations of conditional dependencies that align with domain knowledge about SAG mill operations. This transparency facilitates expert validation, enables incorporation of prior knowledge, and supports diagnostic reasoning when mill performance deviates from expectations [124,125].
3. Uncertainty quantification capabilities vary across methods. Most ML approaches provide point predictions with limited probabilistic interpretation, whereas BNs naturally represent joint probability distributions over all variables [126]. This distinction is crucial for SAG mill applications characterized by significant operations. The ability to propagate uncertainty through the causal network and quantify prediction confidence under partial observability represents a fundamental advantage of the BN framework [127,128].
4. Causal inference and bidirectional reasoning distinguish BNs from alternative ML methods. While XGBoost can predict power consumption given operational parameters, it cannot efficiently infer probable causes of observed power anomalies or evaluate counterfactual scenarios. BNs support these diagnostic and interventional queries through their explicit causal structure, enabling decision support beyond pure prediction [124,129].

The conditions under which ML methods outperform BNs versus vice versa depend on application context. ML methods (particularly deep learning and gradient boosting) demonstrate superior performance when (1) large volumes of high-quality training data are available, (2) the primary objective is accurate point prediction rather than uncertainty characterization, (3) temporal dynamics and sequential dependencies are critical, and (4) computational resources for extensive hyperparameter tuning are accessible [15,68,116]. Conversely, BNs offer advantages when (1) data are limited or incomplete, (2) uncertainty quantification and propagation are essential for risk-informed decision making, (3) causal understanding and diagnostic reasoning are required, and (4) integration of expert knowledge with data is necessary [124,126,130].

The probabilistic nature of BNs aligns naturally with the stochastic variability inherent in SAG mill operations, where ore heterogeneity and equipment wear introduce irreducible uncertainty [105]. While hybrid mechanistic–ML models [115,119] attempt to combine physical interpretability with predictive flexibility, they typically rely on deterministic mechanistic components that may not adequately represent stochastic ore variability. The fully probabilistic BN framework captures both aleatoric uncertainty (inherent randomness in ore properties) and epistemic uncertainty (limited knowledge about system state), providing a more comprehensive representation of operational risk [126,131]. Comparative benchmarking also highlights opportunities for future research integrating BN strengths with ML capabilities. Hybrid architectures combining BN causal structures with ANN or GBM could potentially achieve both high predictive accuracy and rigorous uncertainty quantification. Then, using ML to learn BN structure and parameters from large operational datasets, while constraining structures to respect known causal relationships, represents a promising direction for enhancing performance and interpretability.

To conclude, it is important to emphasize that the proposed BN is a static model and thus does not constitute a complete digital twin. A transition toward a probabilistic digital twin would require explicit temporal modeling, recursive parameter updating, and mechanisms to handle non-stationary ore properties and operational regimes. These components are identified as promising avenues for future research rather than capabilities demonstrated in the present study.

5.3. On Methodological and Practical Recommendations

Some methodological and practical recommendations that should be considered when following future lines of research are presented below:

- Data and sensors: Improved online detection systems and standardized data models. Without reliable signals, probabilistic inference and digital twins cannot be precise [64].
- Combining phenomenology and data: Hybrid models (physics-based ML or surrogate models integrated into simulation) to reduce data requirements and improve extrapolation under uncertainty [64,70].
- Using BNs for decision making and causal analysis: When interpretability and quantified conditional dependencies are important for decision making, BNs are valuable, as they provide probabilistic diagnoses and propagate uncertainty through causal chains [61,62].
- New trend—probabilistic digital twins: A practical roadmap is to build digital twins that incorporate probabilistic models for key components, operate within a simulation system or control loop to evaluate operating modes under uncertainty, and generate optimal control parameters through optimization (evolutionary algorithms) [61,65,68,70].
- Research gaps: Scalable learning of BNs structure with noisy data, rigorous calibration/validation with variable mineralogy, uncertainty propagation for large datasets, and industrial testing with probabilistic controllers [63,64].

Finally, the next practical steps for a probabilistic digital twin of a SAG mill:

- Step 1: Record and track critical variables. Consolidate time series into an event log/record [63,69].
- Step 2: Construct surrogate predictors. Probabilistic BN or hierarchical Bayesian model for causal diagnostics and a high-accuracy ensemble for short-term energy/performance forecasts [15,61,62,68].
- Step 3: Couple the surrogates to a DES/digital twin and a probabilistic optimizer or MPC (Model Predictive Control) that uses uncertainty estimates to compare performance with risk (overhead, energy) [65,68,70].

- Step 4: Run phased industrial pilots with operator validation in the loop and suggestions for conservative setting values before closed-loop control [64].

6. Conclusions and Future Perspectives

6.1. Conclusions

Probabilistic networks have become unique tools for determining and incorporating uncertainty in input variables, due to their ability to provide data for monitoring by incorporating or removing information from their structure, and also to integrate various hypotheses. The performance of classification algorithms based on DBNs presents satisfactory results, as the classification error is low, as validated by goodness-of-fit statistics. Fitting a Discrete Bayesian Network to the SAG mill proved to be a robust approach for capturing key dependencies and delivering explainable inferences. The model generalizes with moderate out-of-sample loss; in test, the accuracy is 0.86 and 0.91 for both SAG Production and SAG Power, indicating useful predictive power in a challenging multiclass problem. When mapping states to an ordinal scale, the $R^2 \sim 0.85$ – 0.90 and $RMSE < 0.5$ bin show that, even when the exact class fails, the error is usually concentrated in adjacent classes, maintaining operational value.

ROC analysis confirmed high separability at the extreme zones (micro-AUC ~ 0.98 ; macro-AUC ~ 0.95 in test, for both SAG Production and SAG Power) and lower discrimination in the intermediate bins, consistent with the superposition of conditions close to the setpoint. Sensitivity and mutual information were consistent with the process physics: SAG Water Feeding and Solid Percentage in the Feeding dominate production (hydraulic/conveying pathway), while rotational speed and Pressure account for a large portion of the SAG Power; particle size distribution plays a modulating role, and liner age has a secondary but stable effect. Overall, the BN offers an interpretable probabilistic framework for diagnosis, prioritization of causes, and “what-if” analysis of control levers.

Finally, a DBN model to analyze SAG production dynamics of the process has the potential to contribute to:

1. Identify the dependencies between the independent variables and the response variable, as well as between the independent variables.
2. Determine the variables that contribute most to explaining the variability of the response(s).
3. Incorporate quantitative knowledge about the frequency of occurrence of an event, using the parameters obtained by the Bayesian network, which will allow for the identification of recurring scenarios.
4. Generate estimates of SAG production based on partial knowledge (in addition to a priori knowledge) of the operational variables considered in the study, such as liner age or mill rotation speed.

The discrete BN’s main value extends beyond predictive accuracy; it provides an interpretable causal structure and a probabilistic mechanism for evaluating operational scenarios under uncertainty. This makes the model a practical decision-support tool for SAG milling, where understanding variable interactions and anticipating system behavior could be more important than exact point predictions.

6.2. Future Perspectives

While the present static BN provides an interpretable probabilistic representation of SAG mill behavior, it is important to highlight that stochastic modeling using BNs applied to the SAG milling process constitutes a first step toward building a robust probabilistic framework, and not yet a fully dynamic formulation. However, the experience gained during the fine-tuning and validation stages suggests multiple avenues for refinement and

expansion that would strengthen the model's representativeness, facilitating its integration into real-world operations. These approaches would not only consider the application of technical improvements, such as discretization, calibration, and sensitivity, but also the evolution toward dynamic architectures, temporal adaptation mechanisms, and the exploration of hybrid approaches.

Discretization based solely on quantile could evolve toward schemes closer to operational reality, incorporating cutoffs defined by physical thresholds such as those associated with pooling, pulp density, or critical speed efficiency. Likewise, the network structure could benefit from greater expert guidance. Strengthening white and black lists, along with additional restrictions on indegree levels, would help reduce structural overfitting and ensure that dependencies more accurately reflect process causality. Along these lines, the extension to dynamic models (such as BN CLG, GBN, or GPR) appears to be a natural evolution, as it would allow capturing deferred temporal effects, such as the impact of the stockpile level on feed PSD and, consequently, on SAG mill response.

Another area of development relates to the model's capability to adapt temporally. Incorporating explicit time-lagged dependencies would enable the model to capture causal sequences that unfold across minutes or hours, which are not represented in the static formulation. Incorporating Bayesian re-estimation strategies using moving windows or online learning schemes would allow monitoring and correcting drift in conditional distributions over campaigns or months, keeping the probabilistic representation up to date with changes in mineralogy or operating conditions. Similarly, exploring multivariate sensitivity analysis would open the possibility of evaluating critical interactions between control levers, such as the combination of solids percentage, water flow, and pressure, under uncertain scenarios.

A Bayesian network with these capabilities would not only allow for the diagnosis of anomalous conditions or the anticipation of deviations, but also for recommending setpoints based on probabilistic optimization criteria, incorporating confidence bands for alarm activation and operational decision making. In this way, the model could evolve from an explainable classifier into a true operational assistant, capable of quantifying the expected impact of each adjustment and supporting process management in highly uncertain environments.

Finally, while the present model provides an interpretable probabilistic representation of SAG mill behavior, it remains a static BN and therefore does not yet fulfill the requirements of a full probabilistic digital twin. Future works should extend this framework toward a dynamic formulation by incorporating temporal dependencies via Dynamic Bayesian Networks (DBNs), modeling time-lagged operational effects, and implementing online parameter updating to address non-stationary process conditions. Then, although the notion of integrating the model into a probabilistic digital twin remains a long-term aspiration, the transition toward a dynamic probabilistic digital twin would require incorporation of temporal dependencies through DBNs or related architectures; implementation of online or incremental updating to reflect evolving process conditions; mechanisms to handle non-stationarity and concept drift; and integration of bidirectional inference into operational decision workflows. These capabilities represent meaningful avenues for future research rather than functionalities demonstrated in the present study.

Author Contributions: Conceptualization, M.S., N.T. and L.A.C.; methodology, M.S. and L.A.C.; software, M.S.; validation, E.G., M.S.-C., E.S.-R., J.C., A.N., D.A. and L.A.C.; formal analysis, M.S. and L.A.C.; investigation, M.S., E.G., M.S.-C., E.S.-R., J.C., A.N., D.A. and L.A.C.; resources, E.G., M.S.-C., E.S.-R. and J.C.; data curation, M.S.; writing—original draft preparation, M.S.; writing—review and editing, M.S.; visualization, M.S.; supervision, A.N., N.T., D.A. and L.A.C.; project administration,

N.T. and L.A.C.; funding acquisition, M.S. and E.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by ANID-Chile, ANID/Fondecyt 1240182.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request from the authors.

Acknowledgments: The authors thank ANID-Chile for funding this research through the ANID/Fondecyt 1240182. Manuel Saldana acknowledges the infrastructure and support of the Doctorado en Ingeniería de Procesos de Minerales of the Universidad de Antofagasta.

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

References

1. Flanagan, D.M. Copper. In *Mineral Commodity Summaries 2025*; U.S. Geological Survey: Reston, VA, USA, 2025; pp. 64–65, ISBN 978-1-4113-4595-9.
2. Velásquez-Yévenes, L.; Torres, D.; Toro, N. Leaching of Chalcopirite Ore Agglomerated with High Chloride Concentration and High Curing Periods. *Hydrometallurgy* **2018**, *181*, 215–220. [\[CrossRef\]](#)
3. Pradhan, N.; Nathsarma, K.C.; Srinivasa Rao, K.; Sukla, L.B.; Mishra, B.K. Heap Bioleaching of Chalcopirite: A Review. *Miner. Eng.* **2008**, *21*, 355–365. [\[CrossRef\]](#)
4. Ghorbani, Y.; Franzidis, J.P.; Petersen, J. Heap Leaching Technology—Current State, Innovations, and Future Directions: A Review. *Miner. Process. Extr. Metall. Rev.* **2016**, *37*, 73–119. [\[CrossRef\]](#)
5. Schlesinger, M.; King, M.; Sole, K.; Davenport, W. *Extractive Metallurgy of Copper*, 5th ed.; Elsevier Ltd.: Amsterdam, The Netherlands, 2011; ISBN 9780080967899.
6. Saldaña, M.; González, J.; Jeldres, R.; Villegas, Á.; Castillo, J.; Quezada, G.; Toro, N. A Stochastic Model Approach for Copper Heap Leaching through Bayesian Networks. *Metals* **2019**, *9*, 1198. [\[CrossRef\]](#)
7. Barros, K.S.; Vielmo, V.S.; Moreno, B.G.; Riveros, G.; Cifuentes, G.; Bernardes, A.M. Chemical Composition Data of the Main Stages of Copper Production from Sulfide Minerals in Chile: A Review to Assist Circular Economy Studies. *Minerals* **2022**, *12*, 250. [\[CrossRef\]](#)
8. Neira, A.; Pizarro, D.; Quezada, V.; Velásquez-Yévenes, L. Pretreatment of Copper Sulphide Ores Prior to Heap Leaching: A Review. *Metals* **2021**, *11*, 1067. [\[CrossRef\]](#)
9. Consejo Minero. *Updated Mining Statistics [Cifras Actualizadas de la Minería]*; Consejo Minero: Santiago, Chile, 2025.
10. Moskalyk, R.R.; Alfanzazi, A.M. Review of Copper Pyrometallurgical Practice: Today and Tomorrow. *Miner. Eng.* **2003**, *16*, 893–919. [\[CrossRef\]](#)
11. Curry, J.A.; Ismay, M.J.L.; Jameson, G.J. Mine Operating Costs and the Potential Impacts of Energy and Grinding. *Miner. Eng.* **2014**, *56*, 70–80. [\[CrossRef\]](#)
12. Bilim, N.; Çelik, A.; Kekeç, B. A Study in Cost Analysis of Aggregate Production as Depending on Drilling and Blasting Design. *J. Afr. Earth Sci.* **2017**, *134*, 564–572. [\[CrossRef\]](#)
13. De Solminihaç, H.; Gonzales, L.E.; Cerda, R. Copper Mining Productivity: Lessons from Chile. *J. Policy Model.* **2018**, *40*, 182–193. [\[CrossRef\]](#)
14. Jeswiet, J.; Szekeres, A. Energy Consumption in Mining Comminution. *Procedia CIRP* **2016**, *48*, 140–145. [\[CrossRef\]](#)
15. Avalos, S.; Kracht, W.; Ortiz, J.M. Machine Learning and Deep Learning Methods in Mining Operations: A Data-Driven SAG Mill Energy Consumption Prediction Application. *Min. Metall. Explor.* **2020**, *37*, 1197–1212. [\[CrossRef\]](#)
16. Daozhen, G.; Peilong, W.; Chunbao, S.; Jue, K.; Ruiyang, Z.; Yang, H.; Daozhen, G.; Peilong, W.; Chunbao, S.; Jue, K.; et al. The Application of JKSimMet Software in the Multi-Objective Collaborative Optimization of the Grinding and Classification System at a Gold Mine. *Conserv. Util. Miner. Resour.* **2020**, *40*, 99–104. [\[CrossRef\]](#)
17. Villanueva, M.; Calderón, C.; Saldaña, M.; Toro, N. Modelling a Sag Grinding System through Multiples Regressions. In Proceedings of the METAL 2020—29th International Conference on Metallurgy and Materials, Brno, Czech Republic, 20–22 May 2020; pp. 1243–1248. [\[CrossRef\]](#)
18. Napier-Munn, J.; Morrell, S.; Morrison, R.D.; Kojovic, T. *Mineral Comminution Circuits Their Operation and Optimization*; Napier-Munn, T.J., Ed.; Julius Kruttschnitt Mineral Research Center: Indooroopilly, Australia, 1996; ISBN 064628861X.
19. Carrasco, C.; Keeney, L.; Napier-Munn, T.J.; Bode, P. Unlocking Additional Value by Optimising Comminution Strategies to Process Grade Engineering® Streams. *Miner. Eng.* **2017**, *103–104*, 2–10. [\[CrossRef\]](#)

20. Morrell, S. Predicting the Overall Specific Energy Requirement of Crushing, High Pressure Grinding Roll and Tumbling Mill Circuits. *Miner. Eng.* **2009**, *22*, 544–549. [[CrossRef](#)]
21. Valery, W., Jr.; Morrell, S.; Kojovic, T.; Kanchibotla, S.; Thornton, D. Modelling and Simulation Techniques Applied for Optimisation of Mine to Mill Operations and Case Studies. In Proceedings of the VI Southern Hemisphere Conference on Minerals Technology, Rio de Janeiro, Brazil, 27 May–1 June 2001; Da Luz, A., Soares, P., Eds.; Corba Editora Artes Graficas: Rio de Janeiro, Brazil, 2001; pp. 107–116.
22. Darwiche, A. *Modeling and Reasoning with Bayesian Networks*, 1st ed.; Cambridge University Press: Los Angeles, CA, USA, 2009; ISBN 9780521884389.
23. McKee, D.J. *Understanding Mine to Mill*, 1st ed.; The Cooperative Research Centre for Optimising Resource Extraction: Brisbane, Australia, 2013; ISBN 9781922029270.
24. Cisternas, L.A.; Lucay, F.A.; Botero, Y.L. Trends in Modeling, Design, and Optimization of Multiphase Systems in Minerals Processing. *Minerals* **2019**, *10*, 22. [[CrossRef](#)]
25. Austin, L.G. A Mill Power Equation for SAG Mills. *Min. Metall. Explor.* **1990**, *7*, 57–63. [[CrossRef](#)]
26. Moys, M.H. A Model of Mill Power as Affected by Mill Speed, Load Volume, and Liner Design. *J. S. Afr. Inst. Min. Metall.* **1993**, *93*, 135–141.
27. Morrell, S. Power Draw of Wet Tumbling Mills and Its Relationship to Charge Dynamics—Part 2: An Empirical Approach to Modelling of Mill Power Draw. *Trans. Inst. Min. Metall. Sect. C Miner. Process. Extr. Metall.* **1996**, *105*, C54–C62.
28. Kojovic, T. Influence of Aggregate Stemming in Blasting on the SAG Mill Performance. *Miner. Eng.* **2005**, *18*, 1398–1404. [[CrossRef](#)]
29. Morrell, S.; Valery, W. Influence of Feed Size on AG/SAG Mill Performance. In Proceedings of the SAG 2001, Vancouver, BC, Canada, 30 September–3 October 2001; pp. 1203–1214.
30. Michaux, S.; Djordjevic, N. Influence of Explosive Energy on the Strength of the Rock Fragments and SAG Mill Throughput. *Miner. Eng.* **2005**, *18*, 439–448. [[CrossRef](#)]
31. Behnamfard, A.; Namaei Roudi, D.; Veglio, F. The Performance Improvement of a Full-Scale Autogenous Mill by Setting the Feed Ore Properties. *J. Clean. Prod.* **2020**, *271*, 122554. [[CrossRef](#)]
32. Morrell, S. The Appropriateness of the Transfer Size in AG and SAG Mill Circuit Design. In Proceedings of the SAG 2011, Los Angeles, CA, USA, 30 January 2011; pp. 1–12.
33. Van Nierop, M.A.; Moys, M.H. Exploration of Mill Power Modelled as Function of Load Behaviour. *Miner. Eng.* **2001**, *14*, 1267–1276. [[CrossRef](#)]
34. Silva, M.; Casali, A. Modelling SAG Milling Power and Specific Energy Consumption Including the Feed Percentage of Intermediate Size Particles. *Miner. Eng.* **2015**, *70*, 156–161. [[CrossRef](#)]
35. Dong, S.; Wang, B.; Wang, Z.; Hu, X.K.; Song, H.C.; Liu, Q. Comparison of Prediction Models for Power Draw in Grinding and Flotation Processes in a Gold Treatment Plant. *J. Chem. Eng. Jpn.* **2018**, *49*, 204–210. [[CrossRef](#)]
36. Lucay, F.A.; Gálvez, E.D.; Salez-Cruz, M.; Cisternas, L.A. Improving Milling Operation Using Uncertainty and Global Sensitivity Analyses. *Miner. Eng.* **2019**, *131*, 249–261. [[CrossRef](#)]
37. Lucay, F.A. Accelerating Global Sensitivity Analysis via Supervised Machine Learning Tools: Case Studies for Mineral Processing Models. *Minerals* **2022**, *12*, 750. [[CrossRef](#)]
38. Li, H.; Evertsson, M.; Lindqvist, M.; Hulthén, E.; Asbjörnsson, G. Dynamic Modeling and Simulation of a SAG Mill-Pebble Crusher Circuit by Controlling Crusher Operational Parameters. *Miner. Eng.* **2018**, *127*, 98–104. [[CrossRef](#)]
39. Asghari, M.; VandGhorbany, O.; Nakhaei, F. Relationship among Operational Parameters, Ore Characteristics, and Product Shape Properties in an Industrial SAG Mill. *Part. Sci. Technol.* **2020**, *38*, 482–493. [[CrossRef](#)]
40. Lvov, V.; Chitalov, L.; Nikolayevna Aleksandrova, T.; Mütze, T. Semi-Autogenous Wet Grinding Modeling with CFD-DEM. *Minerals* **2021**, *11*, 485. [[CrossRef](#)]
41. Marijnissen, M.J.; Graczykowski, C.; Rojek, J. Simulation of the Comminution Process in a High-Speed Rotor Mill Based on the Feed's Macroscopic Material Data. *Miner. Eng.* **2021**, *163*, 106746. [[CrossRef](#)]
42. Ge, Z.; Song, Z.; Ding, S.X.; Huang, B. Data Mining and Analytics in the Process Industry: The Role of Machine Learning. *IEEE Access* **2017**, *5*, 20590–20616. [[CrossRef](#)]
43. Joe Qin, S. Special Issue on Big Data: Data Science for Process Control and Operations. *J. Process Control* **2018**, *67*, iii. [[CrossRef](#)]
44. LV, Y.; Le, Q.T.; Bui, H.B.; Bui, X.N.; Nguyen, H.; Nguyen-Thoi, T.; Dou, J.; Song, X. A Comparative Study of Different Machine Learning Algorithms in Predicting the Content of Ilmenite in Titanium Placer. *Appl. Sci.* **2020**, *10*, 635. [[CrossRef](#)]
45. Tang, J.; Qiao, J.; Liu, Z.; Zhou, X.; Yu, G.; Zhao, J. Mechanism Characteristic Analysis and Soft Measuring Method Review for Ball Mill Load Based on Mechanical Vibration and Acoustic Signals in the Grinding Process. *Miner. Eng.* **2018**, *128*, 294–311. [[CrossRef](#)]
46. Smith, M.L.; Prisbrey, K.A.; Barron, C.L. Blasting Design for Increased SAG Mill Productivity. *Min. Metall. Explor.* **1993**, *10*, 188–190. [[CrossRef](#)]

47. Salazar, J.L.; Magne, L.; Acuña, G.; Cubillos, F. Dynamic Modelling and Simulation of Semi-Autogenous Mills. *Miner. Eng.* **2009**, *22*, 70–77. [\[CrossRef\]](#)
48. Bascur, O.A.; Soudek, A. Grinding and Flotation Optimization Using Operational Intelligence. *Min. Metall. Explor.* **2019**, *36*, 139–149. [\[CrossRef\]](#)
49. Bardinas, J.P.; Aldrich, C.; Napier, L.F.A. Predicting the Operating States of Grinding Circuits by Use of Recurrence Texture Analysis of Time Series Data. *Processes* **2018**, *6*, 17. [\[CrossRef\]](#)
50. Delaney, G.W.; Cleary, P.W.; Morrison, R.D.; Cummins, S.; Loveday, B. Predicting Breakage and the Evolution of Rock Size and Shape Distributions in Ag and SAG Mills Using DEM. *Miner. Eng.* **2013**, *50–51*, 132–139. [\[CrossRef\]](#)
51. Apelt, T.A.; Thornhill, N.F. Inferential Measurement of SAG Mill Parameters V: MPC Simulation. *Miner. Eng.* **2009**, *22*, 1045–1052. [\[CrossRef\]](#)
52. Apelt, T.A.; Thornhill, N.F. Inferential Measurement of Sag Mill Parameters IV: Inferential Model Validation. *Miner. Eng.* **2009**, *22*, 1032–1044. [\[CrossRef\]](#)
53. Bueno, M.P.; Kojovic, T.; Powell, M.S.; Shi, F. Multi-Component AG/SAG Mill Model. *Miner. Eng.* **2013**, *43–44*, 12–21. [\[CrossRef\]](#)
54. Kahraman, A.; Kantardzic, M.; Kahraman, M.M.; Kotan, M. A Data-Driven Multi-Regime Approach for Predicting Energy Consumption. *Energies* **2021**, *14*, 6763. [\[CrossRef\]](#)
55. Olivier, J.; Aldrich, C. Use of Decision Trees for the Development of Decision Support Systems for the Control of Grinding Circuits. *Minerals* **2021**, *11*, 595. [\[CrossRef\]](#)
56. Azizi, A.; Rooki, R.; Mollayi, N. Modeling and Prediction of Wear Rate of Grinding Media in Mineral Processing Industry Using Multiple Kernel Support Vector Machine. *SN Appl. Sci.* **2020**, *2*, 1469. [\[CrossRef\]](#)
57. Hoseinian, F.S.; Abdollahzadeh, A.; Rezai, B. Semi-Autogenous Mill Power Prediction by a Hybrid Neural Genetic Algorithm. *J. Cent. South Univ.* **2018**, *25*, 151–158. [\[CrossRef\]](#)
58. Lu, X.; Kiumarsi, B.; Chai, T.; Jiang, Y.; Lewis, F.L. Operational Control of Mineral Grinding Processes Using Adaptive Dynamic Programming and Reference Governor. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2210–2221. [\[CrossRef\]](#)
59. Xie, Q.; Zhong, C.; Liu, D.; Fu, Q.; Wang, X.; Shen, Z. Operation Analysis of a SAG Mill under Different Conditions Based on DEM and Breakage Energy Method. *Energies* **2020**, *13*, 5247. [\[CrossRef\]](#)
60. Hadizadeh, M.; Farzanegan, A.; Noaparast, M. Supervisory Fuzzy Expert Controller for Sag Mill Grinding Circuits: Sungun Copper Concentrator. *Miner. Process. Extr. Metall. Rev.* **2017**, *38*, 168–179. [\[CrossRef\]](#)
61. Valencia, J.V.; Vargas, F. A Probabilistic Graphical Model for Semi-Autogenous Grinding Processes. In Proceedings of the IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies, ChileCon, Valdivia, Chile, 5–7 December 2023. [\[CrossRef\]](#)
62. Magzumov, Z.; Kumral, M. Application of the Hierarchical Bayesian Models to Analyze Semi-Autogenous Mill Throughput. *Miner. Eng.* **2025**, *232*, 109486. [\[CrossRef\]](#)
63. Pegoraro, M. Probabilistic and Non-Deterministic Event Data in Process Mining: Embedding Uncertainty in Process Analysis Techniques. *CEUR Workshop Proc.* **2022**, *3139*, 37–46. [\[CrossRef\]](#)
64. Estay, H.; Lois-Morales, P.; Montes-Atenas, G.; Ruiz del Solar, J. On the Challenges of Applying Machine Learning in Mineral Processing and Extractive Metallurgy. *Minerals* **2023**, *13*, 788. [\[CrossRef\]](#)
65. Dehon, V.; Quintanilla, P.; Chanona, A.D.R. Probabilistic Model Predictive Control for Mineral Flotation Using Gaussian Processes. *Syst. Control. Trans.* **2025**, *4*, 1023–1028. [\[CrossRef\]](#)
66. Moraga, C.; Astudillo, C.A.; Estay, R.; Maranek, A. Enhancing Comminution Process Modeling in Mineral Processing: A Conjoint Analysis Approach for Implementing Neural Networks with Limited Data. *Mining* **2024**, *4*, 966–982. [\[CrossRef\]](#)
67. Kazemi, M.; Moradkhani, D.; Alipour, A.A. Application of Random Forest and Support Vector Machine for Investigation of Pressure Filtration Performance, a Zinc Plant Filter Cake Modeling. *Int. J. Miner. Process. Extr. Metall.* **2023**, *8*, 15–23. [\[CrossRef\]](#)
68. Ghasemi, Z.; Neshat, M.; Aldrich, C.; Karageorgos, J.; Zanin, M.; Neumann, F.; Chen, L. A Hybrid Intelligent Framework for Maximising SAG Mill Throughput: An Integration of Expert Knowledge, Machine Learning and Evolutionary Algorithms for Parameter Optimisation. *arXiv* **2023**, arXiv:2312.10992. [\[CrossRef\]](#)
69. Nad, A.; Jooshaki, M.; Tuominen, E.; Michaux, S.; Kirpala, A.; Newcomb, J. Digitalization Solutions in the Mineral Processing Industry: The Case of GTK Mintec, Finland. *Minerals* **2022**, *12*, 210. [\[CrossRef\]](#)
70. Quelopana, A.; Ordenes, J.; Wilson, R.; Navarra, A. Technology Upgrade Assessment for Open-Pit Mines through Mine Plan Optimization and Discrete Event Simulation. *Minerals* **2023**, *13*, 642. [\[CrossRef\]](#)
71. Liao, Z.; Xu, C.; Chen, W.; Chen, Q.; Wang, F.; She, J. Effective Throughput Optimization of SAG Milling Process Based on BPNN and Genetic Algorithm. In Proceedings of the 2023 IEEE 6th International Conference on Industrial Cyber-Physical Systems, ICPS 2023, Wuhan, China, 8–11 May 2023. [\[CrossRef\]](#)
72. Bizagi. Bizagi Modeler. 2025. Available online: <https://www.bizagi.com/es/plataforma/modeler> (accessed on 13 July 2025).
73. McCoy, J.T.; Auret, L. Machine Learning Applications in Minerals Processing: A Review. *Miner. Eng.* **2019**, *132*, 95–109. [\[CrossRef\]](#)

74. Umucu, Y.; Deniz, V.; Bozkurt, V.; Fatih Çağlar, M. The Evaluation of Grinding Process Using Artificial Neural Network. *Int. J. Miner. Process* **2016**, *146*, 46–53. [\[CrossRef\]](#)
75. Ahmadzadeh, F.; Lundberg, J. Remaining Useful Life Prediction of Grinding Mill Liners Using an Artificial Neural Network. *Miner. Eng.* **2013**, *53*, 1–8. [\[CrossRef\]](#)
76. Khoshjavan, S.; Khoshjavan, R.; Rezai, B. Evaluation of the Effect of Coal Chemical Properties on the Hardgrove Grindability Index (HGI) of Coal Using Artificial Neural Networks. *J. S. Afr. Inst. Min. Metall.* **2013**, *113*, 505–510.
77. Özbayoğlu, G.; Özbayoğlu, A.M.; Özbayoğlu, M.E. Estimation of Hardgrove Grindability Index of Turkish Coals by Neural Networks. *Int. J. Miner. Process* **2008**, *85*, 93–100. [\[CrossRef\]](#)
78. Venkoba Rao, B.; Gopalakrishna, S.J. Hardgrove Grindability Index Prediction Using Support Vector Regression. *Int. J. Miner. Process* **2009**, *91*, 55–59. [\[CrossRef\]](#)
79. Makokha, A.B.; Moys, M.H. Multivariate Approach to On-Line Prediction of in-Mill Slurry Density and Ball Load Volume Based on Direct Ball and Slurry Sensor Data. *Miner. Eng.* **2012**, *26*, 13–23. [\[CrossRef\]](#)
80. Saldaña, M.; Gálvez, E.; Navarra, A.; Toro, N.; Cisternas, L.A. Optimization of the SAG Grinding Process Using Statistical Analysis and Machine Learning: A Case Study of the Chilean Copper Mining Industry. *Materials* **2023**, *16*, 3220. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Nakhaei, F.; Mosavi, M.R.; Sam, A.; Vaghei, Y. Recovery and Grade Accurate Prediction of Pilot Plant Flotation Column Concentrate: Neural Network and Statistical Techniques. *Int. J. Miner. Process* **2012**, *110–111*, 140–154. [\[CrossRef\]](#)
82. Niedermayer, D. An Introduction to Bayesian Networks and Their Contemporary Applications. In *Innovations in Bayesian Networks*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 156, pp. 117–130, ISBN 978-3-540-85066-3.
83. Koski, T.; Noble, J.M. *Bayesian Networks: An Introduction*; John Wiley: Hoboken, NJ, USA, 2009; ISBN 978-0-470-74304-1.
84. Devore, J. *Probability & Statistics for Engineering and the Sciences*, 8th ed.; Julet, M., Ed.; Cengage Learning: Boston, MA, USA, 2010; ISBN 0-538-73352-7.
85. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2017; ISBN 9780387848570.
86. Grosan, C.; Abraham, A. *Intelligent Systems*, 1st ed.; Grosan, C., Abraham, A., Eds.; Intelligent Systems Reference Library; Springer: Berlin/Heidelberg, Germany, 2011; Volume 17, ISBN 978-3-642-21003-7.
87. Montgomery, D.C.; Runger, G.C. *Applied Statistics and Probabilistly for Engineers*, 6th ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014; ISBN 9781118539712.
88. Pearl, J. *Probabilistic Reasoning in Intelligent Systems*; Elsevier: Amsterdam, The Netherlands, 1988; ISBN 9780080514895.
89. Cooper, G.F.; Herskovits, E. A Bayesian Method for the Induction of Probabilistic Networks from Data. *Mach. Learn.* **1992**, *9*, 309–347. [\[CrossRef\]](#)
90. Heckerman, D.; Geiger, D.; Chickering, D.M. Learning Bayesian Networks: The Combination of Knowledge and Statistical Data. *Mach. Learn.* **1995**, *20*, 197–243. [\[CrossRef\]](#)
91. Tsamardinos, I.; Brown, L.E.; Aliferis, C.F. The Max-Min Hill-Climbing Bayesian Network Structure Learning Algorithm. *Mach. Learn.* **2006**, *65*, 31–78. [\[CrossRef\]](#)
92. Madera, J.; Ochoa, A. Evaluating the Max-Min Hill-Climbing Estimation of Distribution Algorithm on B-Functions. In *Lecture Notes in Computer Science (Including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 11047, pp. 26–33. [\[CrossRef\]](#)
93. Chickering, D.M. Optimal Structure Identification with Greedy Search. *J. Mach. Learn. Res.* **2003**, *3*, 507–554. [\[CrossRef\]](#)
94. Suzuki, J. A Theoretical Analysis of the BDeu Scores in Bayesian Network Structure Learning. *Behaviormetrika* **2017**, *44*, 97–116. [\[CrossRef\]](#)
95. Chen, Y.C.; Wheeler, T.A.; Kochenderfer, M.J. Learning Discrete Bayesian Networks from Continuous Data. *J. Artif. Intell. Res.* **2017**, *59*, 103–132. [\[CrossRef\]](#)
96. Chavira, M.; Darwiche, A.; Jaeger, M. Compiling Relational Bayesian Networks for Exact Inference. *Int. J. Approx. Reason.* **2006**, *42*, 4–20. [\[CrossRef\]](#)
97. Kumral, M. Bed Blending Design Incorporating Multiple Regression Modelling and Genetic Algorithms. *J. S. Afr. Inst. Min. Metall.* **2006**, *106*, 229–237.
98. Ye, Z.; Yahyaei, M.; Hilden, M.; Powell, M.S. Novel Size Segregation Indices for Multi-Sized Particle Stockpiles. *Miner. Eng.* **2023**, *201*, 108165. [\[CrossRef\]](#)
99. Saavedra, M.; Risso, N.; Momayez, M.; Nunes, R.; Tenorio, V.; Zhang, J. Blending Characterization for Effective Management in Mining Operations. *Minerals* **2025**, *15*, 891. [\[CrossRef\]](#)
100. Li, H.; Asbjörnsson, G.; Bhadani, K.; Evertsson, M. Investigating Dynamic Behavior in SAG Mill Pebble Recycling Circuits: A Simulation Approach. *Minerals* **2024**, *14*, 716. [\[CrossRef\]](#)
101. Liu, Y.; Spencer, S. Dynamic Simulation of Grinding Circuits. *Miner. Eng.* **2004**, *17*, 1189–1198. [\[CrossRef\]](#)

102. Powell, M.S.; Valery, W. Slurry Pooling and Transport Issues in SAG Mills. In Proceedings of the International Autogenous and Semiautogenous Grinding Technology, Vancouver, BC, Canada, 24–27 September 2006; University of British Columbia, Dept. of Mining Engineering: Vancouver, BC, Canada, 2006; Volume 1, pp. 133–152.
103. Mulenga, F.K.; Moys, M.H. Effects of Slurry Pool Volume on Milling Efficiency. *Powder Technol.* **2014**, *256*, 428–435. [\[CrossRef\]](#)
104. Guo, W.; Guo, K. Effect of Solid Concentration on Particle Size Distribution and Grinding Kinetics in Stirred Mills. *Minerals* **2024**, *14*, 720. [\[CrossRef\]](#)
105. Weerasekara, N.S.; Powell, M.S. Performance Characterisation of AG/SAG Mill Pulp Lifters Using CFD Techniques. *Miner. Eng.* **2014**, *63*, 118–124. [\[CrossRef\]](#)
106. Toor, P.; Powell, M.; Hilden, M.; Weerasekara, N. Understanding the Effects of Liner Wear on SAG Mill Performance. In Proceedings of the MetPlant 2015, Perth, Australia, 7 September 2015; The Australasian Institute of Mining and Metallurgy: Perth, Australia, 2015; pp. 150–161.
107. Saldaña, M.; Neira, P.; Flores, V.; Robles, P.; Moraga, C. A Decision Support System for Changes in Operation Modes of the Copper Heap Leaching Process. *Metals* **2021**, *11*, 1025. [\[CrossRef\]](#)
108. Rogers, P. *Advances in Computational Intelligence Applications in the Mining Industry*; MDPI: Basel, Switzerland, 2022; ISBN 978-3-0365-3158-8.
109. Ghasemi, Z.; Neumann, F.; Zanin, M.; Karageorgos, J.; Chen, L. A Comparative Study of Prediction Methods for Semi-Autogenous Grinding Mill Throughput. *Miner. Eng.* **2024**, *205*, 108458. [\[CrossRef\]](#)
110. Pural, Y.E.; Ledezma, T.; Hilden, M.; Forbes, G.; Boyle, F.; Yahyaei, M. Application of Machine Learning for Generic Mill Liner Wear Prediction in Semi-Autogenous Grinding (SAG) Mills. *Minerals* **2024**, *14*, 1200. [\[CrossRef\]](#)
111. Ghasemi, Z.; Neshat, M.; Aldrich, C.; Zanin, M.; Chen, L. Optimising SAG Mill Throughput and Circulating Load Using Machine Learning Models: A Multi-Objective Approach for Identifying Optimal Process Parameters. *Miner. Eng.* **2025**, *232*, 109551. [\[CrossRef\]](#)
112. Ruiz, M.A.V.; Gonzales, J.A.V.; Villalba, F.J.B. Multivariable Predictive Models for the Estimation of Power Consumption (KW) of a Semi-Autogenous Mill Applying Machine Learning Algorithms [Modelos Predictivos Multivariables Para La Estimación de Consumo de Potencia (KW) de Un Molino Semi—Autógeno aplicando algoritmos de Machine Learning]. *J. Energy Environ. Sci.* **2024**, *8*, 14–31. [\[CrossRef\]](#)
113. Meinshausen, N. Quantile Regression Forests. *J. Mach. Learn. Res.* **2006**, *7*, 983–999.
114. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 13–17 August 2016; pp. 785–794. [\[CrossRef\]](#)
115. Feng, Y.; Wang, X.; Zou, H.; Yan, L. A Composite Power Prediction Model for Semi-Au-Togenous Grinding Mill Based on Mechanistic Approach and XGBoost. In Proceedings of the 37th Chinese Control and Decision Conference, CCDC, Xiamen, China, 16–19 May 2025; pp. 557–564. [\[CrossRef\]](#)
116. Lopez, P.; Reyes, I.; Risso, N.; Aguilera, C.; Campos, P.G.; Momayez, M.; Contreras, D. Assessing Machine Learning and Deep Learning-Based Approaches for SAG Mill Energy Consumption. In Proceedings of the 2021 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies, CHILECON, Valparaíso, Chile, 6–9 December 2021. [\[CrossRef\]](#)
117. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#) [\[PubMed\]](#)
118. Zhang, D.; Xiong, X.; Shao, C.; Zeng, Y.; Ma, J. Semi-Autogenous Mill Power Consumption Prediction Based on CACN-LSTM. *Appl. Sci.* **2024**, *15*, 2. [\[CrossRef\]](#)
119. Hermosilla, R.; Valle, C.; Allende, H.; Aguilar, C.; Lucic, E. SAG's Overload Forecasting Using a CNN Physical Informed Approach. *Appl. Sci.* **2024**, *14*, 11686. [\[CrossRef\]](#)
120. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#)
121. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; The MIT Press: Cambridge, MA, USA, 2016; ISBN 9780262337373.
122. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased Boosting with Categorical Features. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, QC, Canada, 2–8 December 2018. [\[CrossRef\]](#)
123. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000. [\[CrossRef\]](#)
124. Pearl, J. *Causality: Models, Reasoning, and Inference*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2011; pp. 1–464. [\[CrossRef\]](#)
125. Koller, D.; Friedman, N. Structure Learning in Bayesian Networks. In *Probabilistic Graphical Models: Principles and Techniques*; The MIT Press: Cambridge, MA, USA, 2009; pp. 783–848.
126. Pourret, O.; Naim, P.; Marcot, B. Introduction to Bayesian Networks. In *Bayesian Networks: A Practical Guide to Applications*; Wiley: Hoboken, NJ, USA, 2008; pp. 1–430. [\[CrossRef\]](#)

127. Fenton, N.; Neil, M. *Risk Assessment and Decision Analysis with Bayesian Networks*, 2nd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2018; ISBN 9781315269405.
128. Kjærulff, U.B.; Madsen, A.L. *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 22. [[CrossRef](#)]
129. Bareinboim, E.; Pearl, J. Causal Inference and the Data-Fusion Problem. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 7345–7352. [[CrossRef](#)]
130. Marcot, B.G.; Penman, T.D. Advances in Bayesian Network Modelling: Integration of Modelling Technologies. *Environ. Model. Softw.* **2019**, *111*, 386–393. [[CrossRef](#)]
131. Der Kiureghian, A.; Ditlevsen, O. Aleatory or Epistemic? Does It Matter? *Struct. Saf.* **2009**, *31*, 105–112. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.