-RESEARCH PAPER-

International Journal of Mining and Geo-Engineering

IJMGE

IJMGE 59-1 (2025) 83-89

DOI: 10.22059/IJMGE.2025.383890.595204

Improving energy efficiency in the mining industry: an LSTM-ANN predictive model for sieve refusal in grinding mills

Chaimae Loudari [a, *], Moha Cherkaoui [a], Imad El Harraki [a], Rachid Bennani [b], Mohamed El Adnani [c], EL Hassan Abdelwahed [c], Intissar Benzakour [d], François Bourzeix [e] and Karim Baina[f]

[a] *LMAID Lab, National School of Mines of Rabat, ENSMR, Rabat, Morocco.*

[b] *Digitalization and Microelectronic Smart Devices, MAScIR, Rabat, Morocco.*

[c] *LISI Laboratory University of Cadi Ayyad, UCA, Marrakech, Morocco.*

[d] *Reminex Research Center, MANAGEM Group, Marrakech, Morocco.*

[e] *Embedded Systems and Artificial Intelligence Department, MAScIR, Rabat, Morocco.*

[f] *Al-Qualsadi Research and Development Team, ENSIAS, Rabat, Morocco.*

Article History:

Received: 28 October 2024.

Revised: 11 December 2024.

Accepted: 12 February 2025.

ABSTRACT

The integration of Artificial Intelligence (AI) within Mine 4.0 has significantly advanced the mining industry by enhancing performance, efficiency, safety, and overall productivity. Despite these advancements, a critical gap exists in predicting sieve refusal, a key parameter affecting grinding mill efficiency and product quality, particularly when accounting for the temporal and nonlinear dependencies inherent in mining data. This study introduces a hybrid predictive model that combines Long Short-Term Memory networks (LSTM) with Artificial Neural Networks (ANN) to predict sieve refusal in grinding mills utilizing actual process and energy data from mining industry databases. The LSTM component captured the temporal dynamics and time-delayed dependencies of variables and power consumption. Concurrently, the ANN component modeled complex, nonlinear relationships among input variables. This hybrid approach effectively addressed the intrinsic characteristics of mining data, which are often overlooked in traditional models. Comparative analyses demonstrated that the proposed LSTM-ANN model significantly outperformed existing advanced regression and deep learning methods, establishing it as a state-of-the-art solution for sieve refusal prediction. The enhanced predictive accuracy provided direct support for operational planning and scheduling, contributing to improved energy efficiency and cost-effectiveness in mining operations. By addressing the underexplored temporal and spatial interrelations between variables and sieve refusal, this research fills a notable gap in the application of deep learning to grinding mill operations. The findings underscore the transformative potential of advanced AI models in optimizing mining practices, aligning with the broader objectives of Mine 4.0 to leverage intelligent data analysis for operational excellence.

Keywords: *Energy eficiency, Grinding mills, LSTM-ANN model, Mining industry, Sieve refusal prediction.*

1. Introduction

The transition to Mine 4.0 marks a transformative era in the mining industry, integrating artificial intelligence (AI) [1] [2] to enhance performance, efficiency, safety, and productivity. AI-driven technologies [3-6] have significantly reduced production costs and energy consumption [7], reflecting substantial advancements in the sector [8][9]. Central to Mine 4.0 is the use of large datasets [10], enabled by the Industrial Internet of Things (IIoT), cloud storage, and interactive dashboards. Machine learning algorithms [11] optimize parameters, boosting efficiency and reducing human intervention [12]. These technologies highlight Mine 4.0's transformative impact on mining practices [13-15].

Energy efficiency remains a critical focus, as the mining sector accounts for 11% of global energy usage, 38% of industrial energy consumption, and 15% of global electricity use [16]. Grinding mills, responsible for ore size reduction, are particularly energy-intensive due to operational inefficiencies [17]. Emerging regression techniques in machine learning leverage real-time data for predictive analysis [18][19], offering an alternative to traditional models. Advanced models, such as support vector regression (SVR), polynomial regression (PR), and ensemble methods, including Random Forest and XG-Boost, provide superior accuracy in predicting grinding mill energy needs [20-22].

Deep learning methods, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, have shown high accuracy in predicting energy consumption for grinding mills [23][24]. However, many studies have failed to consider the temporal and spatial interrelations between variables, which were critical for accurately predicting sieve refusal. Addressing this gap, the proposed Long Short-Term Memory-Artificial Neural Network (LSTM-ANN) model was specifically designed to predict sieve refusal in grinding mills, offering

* Corresponding author. *E-mail address:* chaimae.loudari@enim.ac.ma (C. Loudari).

several distinct advantages over existing approaches. By leveraging LSTM's ability to handle time-dependent variables and ANN's capacity to model complex nonlinear interactions, the LSTM-ANN outperformed traditional methods, such as Random Forest, XG-Boost, and standalone GRU models. Validated using real mining industry data, the model demonstrated practical applicability, achieving superior predictive accuracy, while supporting optimized scheduling, reduced energy consumption, and cost-effective operations. These strengths positioned the LSTM-ANN as a robust and reliable solution for improving energy efficiency in grinding mill processes.

The main contributions of this article are as follows:

- The development and validation of an LSTM-ANN model using real mining industry data.
- The application of LSTM to capture temporal dynamics and ANN to model nonlinear critical information.
- The demonstration of the LSTM-ANN model's superiority over existing methods.
- The provision of precise predictive insights for operational planning, enhancing energy efficiency and cost-effectiveness.

The paper is structured as follows: Section 2 identifies key variables influencing sieve refusal and explores data characteristics. Section 3 outlines the proposed predictive model. Section 4 analyzes model performance, and Section 5 concludes with findings and implications.

2. Material and methodology

2.1. Rod grinding mills

In the CMG plant, a Moroccan mining plant that produces lead, copper, and zinc ores, the ores are crushed before being fed with water into the rod grinding mills, which play a critical role as energy-intensive equipment [25-27]. These mills primarily function to reduce the size of ore particles from D80 8mm to D80 500µm through impact and abrasion caused by rod grinding charges, typically in suspension within water, by adopting continuous feeding and continuous rolling mode. The grinding process employed is an overflow grinding system, with a hydro-cyclone used to separate the overflow and underflow based on their D80 particle sizes.

Ideally, the overflow fraction should have a 20% rejection rate, as determined by a specific sieve size. If the rejection rate exceeds 20%, it implies that the smaller mineral particles need to be returned to the grinding process for regrinding, which increases the overall energy consumption substantially. Conversely, if the rejection rate falls below 20%, it indicates that the mineral particles are smaller than the desired D80 size. This suggests that the grinding intensity is higher than necessary, resulting in increased energy consumption during the grinding process. Thus, maintaining a refusal rate of precisely 20% during grinding ensures optimal performance and energy efficiency.

2.2. Dataset

Based on comprehensive research and field visits conducted at the CMG plant, several variables have been identified as influential factors in the energy consumption of grinding mills. These variables include the feed tonnage and feed water to the grinder, the power consumption during the grinding process, the sieve refusal of the mill's discharge, and the proportion of minerals from each mine source entering the grinder.

The variables were collected from the operational data of rod grinding mills at the CMG plant, with varying frequencies based on their utilization in the process operation. The feed tonnage (t/h) and feed water (m3/h) were collected at a rate of 30 observations per minute, while the power (kW) was recorded at 120 observations per minute. These three variables were collected using sensors specifically installed in the rod grinding mills.

Data collection for sieve refusal (% of particle size +500 µm) was performed manually with only three observations obtained per day. Samples were taken directly from the grinder and sent to the laboratory. There, they were analyzed using sieve screens to determine the proportion of particles exceeding a specific size threshold.

There were three distinct mine sources providing minerals for the ore feeding the grinder. For each shift throughout the day, which consists of three shifts, mining experts examined the tonnage proportion of each mine source. These shifts were designated as follows: shift 1 from 7 a.m. to 3 p.m., shift 2 from 3 p.m. to 11 p.m., and shift 3 from 11 p.m. to 7 a.m. Table 1 presents the summary statistics for the dataset.

2.3. Preprocessing dataset

In order to build the dataset, it is necessary to unify the frequency of all the variables. Therefore, the frequency of all variables is reduced to three observations per day due to the low frequency of sieving rejects in grinders . Frequency reduction for the other variables is accomplished by calculating the average of each variable during each shift.

Then all variables are combined into a single dataset, in which each row contains a NaN value or one or more columns have a value of 0 has been eliminated. As a result, the data set consists of 256659 observations. Table 1 presents the summary statistics for the dataset.

The purpose of this pre-processing step is to generate a dataset that accurately represents the operational characteristics of the grinding mills.

Table 1. The descriptive statistics of the process variables.

Variables	Min	Max	Mean	Stdv
Feed Water (m^3/h)	1.28	62.63	10.53	1.49
Feed Tonnage (T/h)	1.88	181.91	118.32	13.33
Sieve Refusal (%)	12.69	54.21	24.91	5.45
Power (KW)	8.22	718.69	508.57	42.00
Ore 1 (%)	0	100	81.65	17.56
Ore 2 (%)	0	100	14.92	8.93
Ore 3 (%)	0	100	3.43	18.20

2.4. Correlation matrix

The correlation matrix Fig. 1 was applied to the dataset variables in order to analyze the degree of correlation between these variables. The degree of correlation between two variables ranged from -1 to 1. A positive correlation between two variables indicated that the correlation coefficient was above 0, with higher values signifying a stronger positive correlation. A negative correlation between two variables meant that the correlation coefficient was below 0, and lower values suggested a stronger negative correlation. No correlation between two variables implied that the correlation coefficient was 0. The aim of this study was to predict the sieve refusal of the mill's discharge. The correlation coefficients between this output and the various input variables could be either positive or negative, and this was justified by the relationships between the sieve refusal and each of these other variables. A positive correlation existed between the sieve refusal and both feed tonnage and feed water. This was because an increase in these feeds led to an increase in sieve refusal, as the grinder mill did not reduce the particles to a suitable size. Conversely, there was a negative correlation between sieve refusal and grinder power. This was because when the grinder consumed more power, it more effectively reduced the particle size, leading to a decrease in sieve refusal. The correlation coefficients between the sieve refusal and the proportion of minerals from each mining source depended on the unique characteristics of each mineral. These findings suggested that the selected input variables were appropriate for predicting the output.

2.5. Performance metric

The evaluation of model performance utilized selected performance metrics, including (determination of coefficient) R^2, (Mean Square Error) MSE, (Mean Absolute Error) MAE, and (Mean Absolute Percentage Error) MAPE. R^2 quantifies the degree of correlation between predicted and true values, serving as a measure of the model's goodness of fit [28]. Eq. (1) delineated the computation for R^2. A higher R^2 value signifies a more accurate fit to the data.

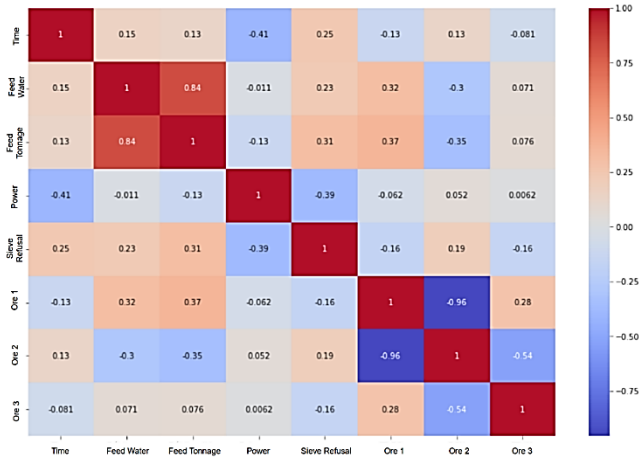


Fig. 1. The correlation matrix.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (1)$$

Where $\hat{y}_i$ denotes the predicted values, y_i represents the actual values, n is the total number of samples, and $\bar{y}_i$represents the mean sieve refusal from the sample.

The MSE quantifies the average squared discrepancy between the predicted values and the actual values [29]. Eq. (2) provides the formula for computing the MSE.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \quad (2)$$

The MAE signifies the mean of all absolute differences, reflecting the average discrepancy between predicted and actual values [30]. Eq. (3) outlines the computation for the MAE.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \quad (3)$$

3. Proposed method

This paper proposed a hybrid LSTM-ANN model designed for the prediction of sieve refusal in grinding mills, targeting the optimization of grinding efficiency and product quality. The hybrid approach employed the LSTM model to adeptly recognize sequential and temporal dependencies within the dataset's features. Concurrently, the ANN model was utilized to discern intricate non-linear associations present in the same dataset. The distinction from traditional methods was that in our approach, grinding mills were complex, large systems with significant inertia and extended delays. Since the historical data of all variables served as crucial references for prediction tasks, previous variable data were incorporated into the LSTM-ANN model as part of its input. Fig. 2 shows the structure of the LSTM-ANN.

3.1. Long Short-Term Memory (LSTM) architecture

The LSTM model is a type of recurrent neural network (RNN) developed to capture long-term dependencies in time-series data. Its structure is based on three main gates: the forget gate Eq. (4), the input gate Eq. (5) and Eq. (6), the update of cell state Eq. (7), and the output gate Eq. (8) and Eq. (9), which together control the flow of data through the cell. Given an input sequence $x = (x_1, x_2, ..., x_t)$.

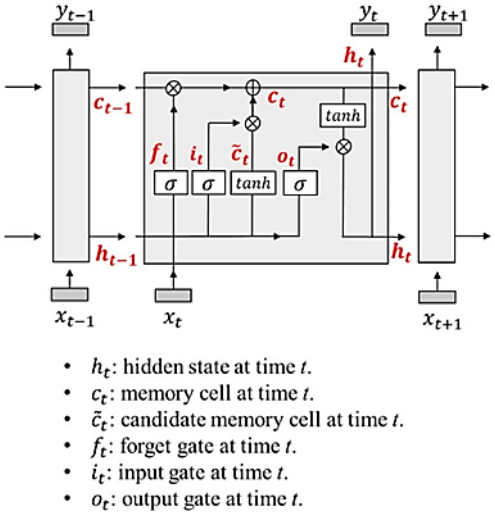
Forget gate:

$$f_t = \sigma\,(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

Input gate:

$$i_t = \sigma\,(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

$$\tilde{C}t = tanh\,(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (6)$$



- h_t: hidden state at time t.
- c_t: memory cell at time t.
- $\tilde{c}_t$: candidate memory cell at time t.
- f_t: forget gate at time t.
- i_t: input gate at time t.
- o_t: output gate at time t.

Fig. 2. The structure of the LSTM model [21].

Update of cell state:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}t \quad (7)$$

Output gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = o_t \times \tanh(C_t) \quad (9)$$

Where $[h_{t-1}, x_t]$ denotes the concatenation of the previous hidden state h_{t-1} and the current input x_t. The symbol σ refers to the sigmoid function, serving as an activation function within the model. This is particularly important for controlling the gates in the LSTM cell. Similarly, tanh(x) is the hyperbolic tangent function, another key activation function used in the LSTM structure for modulating the cell state. The matrices W_i, W_o, W_f, *and* W_c represent the weights, and b_i, b_o, b_f, *and* b_c represent the biases associated with the input, output, forget, and cell state gates, respectively. These weights and biases are crucial parameters that the model learns through the training process.

The forget gate f_t evaluates the significance of the information in x_t to determine how much of the past data should be discarded. Conversely, the input gate i_t assesses x_t to decide what new information should be absorbed. Both f_t and i_t utilize the sigmoid function for activation, applied to a linear combination of x_t and h_{t-1}. A prospective memory cell $\tilde{c}_t$ is then created by processing this linear combination through a Tanh function. The updated memory cell c_t is formed by adding two components: the portion of the previous memory cell c_{t-1} to be retained, found by element-wise multiplication with f_t, and the new information from the candidate memory cell, deter- mined by element-wise multiplication of i_t and $\tilde{c}_t$. The output gate o_t derives from a linear combination of x_t and h_{t-1} passed through a sigmoid function. This gate regulates how much information from the current memory cell c_t is transferred to the final hidden state h_t, calculated through an element-wise multiplication of o_t and the Tanh of c_t.

3.2. Artificial Neural Network (ANN) architecture

After processing through the LSTM layers, the extracted sequence features were input into an ANN model. The purpose of the ANN was to discern complex and non-linear relationships between variables to enhance prediction accuracy. The output h from the LSTM served as the input for the ANN. The final prediction result is then presented as per the equation denoted by Eq. (10).

$$\hat{y} = W_{ann}h + b_{ann} \quad (10)$$

Where $\hat{y}$ represents the predicted output, W_{ann} and b_{ann} denote the weights and biases of the ANN, respectively.

3.3. Prediction

The hybrid LSTM-ANN model connects the output of the LSTM to the ANN layer, combining the ability of the LSTM to capture time-dependent features with the capacity of the ANN for non-linear combinations of features.

The overall equation representing the output prediction results of the hybrid model is:

$$\dot{y} = W_{ann}(o_t \times \tanh(C_t)) + b_{ann} \quad (11)$$

Through iterative training, the hybrid model refines its weights and biases, converging towards a result that minimizes the prediction error between the actual and predicted sieve refusal values.

3.4. Optimization, training and hyperparameters

Experimentally, the TensorFlow framework was used to build the hybrid LSTM-ANN model. To achieve optimal performance, accurate tuning of the model's hyperparameters was essential. The dataset was structured as a 2D array with dimensions NxM, where N = 256,659 observations and M = 7 features: Feed Water (m³/h), Feed Tonnage (T/h), Power (KW), Ore 1 (%), Ore 2 (%), Ore 3 (%), and a time variable. These features were used to predict the target output variable, Sieve Refusal (%), which represents the proportion of particles exceeding a specific size threshold. The architecture of the hybrid LSTM-ANN model and the selected hyper-parameters are summarized in Table 2. The learning rate was set to 0.0001, chosen after a series of experiments to balance the speed of convergence and stability. The model was trained for 1204 epochs using the Adam optimizer, which adaptively adjusts learning rates for each parameter, ensuring efficient convergence and strong performance. Additionally, the batch size was set to 128 to strike a balance between computational efficiency and model accuracy. Early stopping criteria were applied to avoid overfitting by monitoring the validation loss during training.

Table 2. Hyperparameter Configuration of the LSTM-ANN Model.

Hyperparameter	Value
Num-LSTM-Layers	5
Num-Dense-Layers	8
Units-LSTM-1	144
Units-LSTM-2	48
Units-LSTM-3	160
Units-LSTM-4	256
Units-LSTM-5	256
Units-Dense-1	128
Units-Dense-2	256
Units-Dense-3	16
Units-Dense-4	224
Units-Dense-5	208
Units-Dense-6	224
Units-Dense-7	192
Units-Dense-8	1
Batch Size	128
Learning Rate	0.0001
Optimizer	Adam

4. Results and discussions

In order to demonstrate the superior predictive capabilities of our proposed prediction model in estimating sieve refusal, we compared it with several other models, including MLP, ANN, and GRU-ANN. We then analyzed the performance of these different models on our datasets using metrics, such as R^2, MSE, and MAE.

4.1. Sieve refusal prediction

The LSTM-ANN model proposed in this paper effectively addresses issues related to delay and methodological complexity, enabling the accurate prediction of sieve refusal in grinding mills. In this section, we compared the predictive performance of various models for sieve refusal in grinding mills. Experimentally, the input comprises six variables, with the prediction target being the percentage of sieve refusal. Fig. 3 illustrates the actual sieve refusal values with the blue line and the predicted values with the red line.

In order to validate the superiority of the LSTM-ANN, we conducted a comparison with various models. The Multi-Layer Perceptron (MLP) was employed for diverse machine learning tasks, including classification and regression. The ANN was utilized in tasks, such as pattern recognition and decision-making. On the other hand, the GRU-ANN is a hybrid model that integrates the Gated Recurrent Unit (GRU), a type of recurrent neural network, with conventional ANN components. This hybrid design is tailored for effectively capturing sequential data dependencies and is commonly applied in tasks involving time-series data, natural language processing, and more. In terms of prediction results, the deep learning approach produced distinct outcomes for this dataset. Fig. 3 illustrates the performance of an MLP model, which exhibits a basic level of predictive capability. While it broadly mirrored the general direction of actual values, it fell short in capturing their complexity, especially at the peaks and troughs. This indicated a fundamental understanding of the data, suitable for recognizing broad trends but lacking in precision for detailed nuances. The ANN model displayed in the graph showed a slight improvement over the MLP. It appeared to better grasp some of the abrupt changes in the data, yet it still struggles with finer details, as evident from the missed peaks. This suggested a model with a basic temporal understanding but without the sophistication required for precise predictions. Moving to the GRU-ANN, we observed a significant improvement in predictive accuracy. The graph demonstrated the model's ability to closely follow the data's movements, including some of the sharper turns. The gated mechanisms of the GRU-ANN, which facilitated effective information retention and forgetting, contributed to this enhanced capability. Finally, the LSTM-ANN graph presented the most accurate predictions. It not only captured the general trend of actual values but also intricated patterns, including sharp spikes. The exceptional performance of the LSTM-ANN could be attributed to its long-term memory cells, which excelled at retaining information over extended periods, facilitating a nuanced understanding of the data. In summary, while all models exhibited varying degrees of predictive ability in capturing the deployment sample's trend, the LSTM-ANN model clearly stood out. Its capacity to precisely match the actual data across both high and low points indicated a depth of learning that other models do not achieve. For a deployment sample characterized by a rich tapestry of fluctuations, the LSTM-ANN model's nuanced approach to data interpretation established it as the superior choice for accurate predictions. The LSTM-ANN model outperformed other methods due to its ability to effectively combine long-term memory and nonlinear feature learning. Unlike the MLP and ANN models, which lacked the capacity to capture temporal dependencies and complex interactions in the data, the LSTM-ANN leveraged its memory cells to retain critical sequential information. The GRU-ANN model, while capable of modelling some time-series relationships, fell short in handling extended dependencies, limiting its predictive accuracy. In contrast, the LSTM-ANN excelled at both capturing intricate patterns and minimizing prediction errors, as demonstrated by its superior performance metrics, including the highest R^2 value and the lowest MSE and MAE. This hybrid architecture enabled the LSTM-ANN to achieve unparalleled consistency and reliability, making it the most robust choice for predicting sieve refusal in grinding mills.

4.2. Model evaluation

A concise and informative comparison of four different neural network algorithms is presented in Table 3, using standard performance evaluation metrics commonly applied in the field. The R^2 values reflect each model's ability to capture the variance of the observed data, with higher values indicating a better fit. Here, the LSTM-ANN model scored the highest with an R^2 of 0.89, suggesting that it can

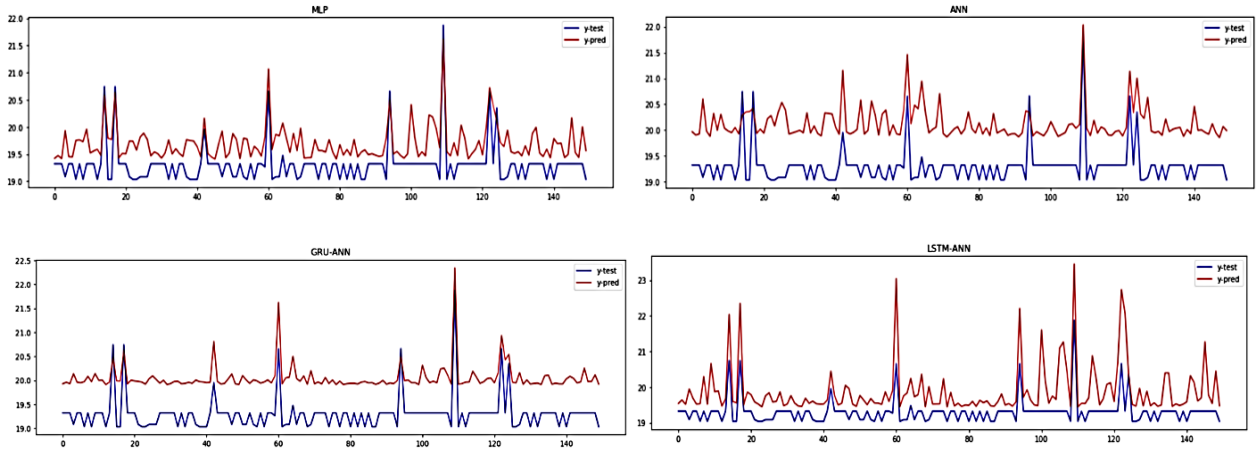


Fig. 3. The distribution of prediction values and true values.

predict 89% of the variance, which is a substantial proportion, indicating a strong predictive performance. The MSE and MAE offered perspectives on the average magnitude of the models' prediction errors, with lower values being preferable. Again, the LSTM-ANN model demonstrated its superiority with the lowest MSE and MAE, signifying that its predictions were closest to the actual values on average. In contrast, the MLP exhibited the weakest fit with the highest error metrics, whereas the ANN and GRU-ANN hold intermediate positions, with the GRU-ANN closer to the LSTM-ANN but not surpassing it. This data unequivocally pointed to the LSTM-ANN as the most precise and reliable model among those tested, making it the standout choice for tasks requiring high accuracy in predictive modelling.

Table 3. Model's performances.

Algorithm	R2	MSE	MAE
MLP	0.57	5.72	1.70
ANN	0.77	2.22	1.11
GRU-ANN	0.85	1.43	0.89
LSTM-ANN	0.89	1.08	0.77

4.3. Prediction values distribution

The comparative analysis of predictive models, as illustrated in Fig. 4, highlighted key insights into their predictive consistency and bias by examining the mean residuals and standard deviations of residuals. The MLP model, with mean residuals near zero (0.024), suggested minimal bias in predictions; however, the relatively high standard deviation of residuals (2.203) indicated a lack of consistency in the predictive performance across the dataset. The ANN model reduced the variability in predictions, as evidenced by a lower standard deviation of residuals (1.473), but exhibited a slight negative bias with mean residuals at -0.232. The GRU-ANN model further improved upon consistency with a standard deviation of residuals at 1.163, and a mean residual of -0.279, which indicated a modest negative bias but with more reliable prediction outcomes compared to the MLP and ANN models. Notably, the LSTM-ANN model outperformed all others by demonstrating the highest accuracy and the most consistent results, with the lowest standard deviation of residuals (1.027) and mean residuals of -0.182, suggesting that while there was a slight underestimation in the predictions, they were reliably close to the true values across different instances. The analysis distinctly positioned the LSTM-ANN model as the most robust and reliable for modelling and forecasting tasks within our examined dataset, highlighting its potential for deployment in scenarios demanding high precision.

The LSTM-ANN model outperformed other methods at all times in making predictions, with GRU-ANN and ANN coming in next. This means that LSTM-ANN was best at understanding the data it was given. It can reliably and correctly forecast how much material a grinding mill's sieve will reject. This proved that the LSTM-ANN is a good choice for predicting sieve performance in the mining industry.

4.4. Comparative analysis of prediction errors

The comparative analysis of error distributions between the actual and predicted values for different models is presented in Fig. 5. These figures illustrate the errors for the MLP, ANN, GRU-ANN, and the proposed LSTM-ANN models. The results showed that the LSTM-ANN model achieved the most consistent and minimal prediction errors compared to the other models. Specifically, the LSTM-ANN exhibited reduced error magnitudes and fewer extreme deviations, indicating its superior ability to learn and capture both sequential dependencies and complex nonlinear relationships in the data. In contrast, the MLP and ANN models struggled with larger error variations, particularly in regions with high fluctuations, reflecting their limitations in handling time-series data. The GRU-ANN model, while performing better than MLP and ANN, still fell short of the precision achieved by the LSTM-ANN. This analysis underscored the LSTM-ANN's robustness and reliability, making it the most suitable model for accurately predicting sieve refusal in grinding mill systems.

5. Conclusions

This research proposed a novel LSTM-ANN hybrid model for predicting sieve refusal in grinding mill systems. By combining the temporal pattern recognition capabilities of LSTM with the nonlinear relationship modelling of ANN, the model delivered accurate and stable predictions. Empirical validation using a real-world dataset demonstrated its superiority over other models, such as MLP, ANN, and GRU-ANN, with performance metrics, including an MAE of 0.77, an MSE of 1.08, and an R^2 of 0.89. These results highlight the LSTM-ANN model's potential to enhance predictive accuracy in grinding mill operations within the mining industry.

While effective, the model's extensive parameter set imposes computational challenges, resulting in slower training speeds. Future research will focus on optimizing the model for computational efficiency, aiming to develop a lightweight version that maintains predictive performance while improving applicability in resource-constrained environments. This work lays a foundation for further advancements in predictive modelling and energy optimization in mining operations.

Declaration of competing interest

The authors declare that they have no known competing financial

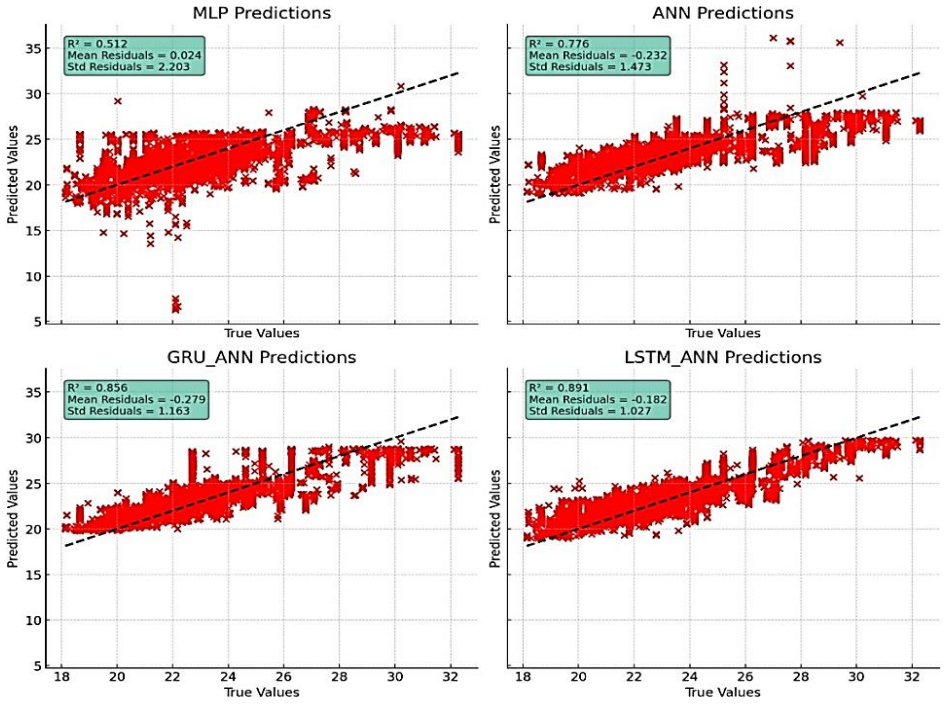


Fig. 4. The comparative analysis of the predictive models.

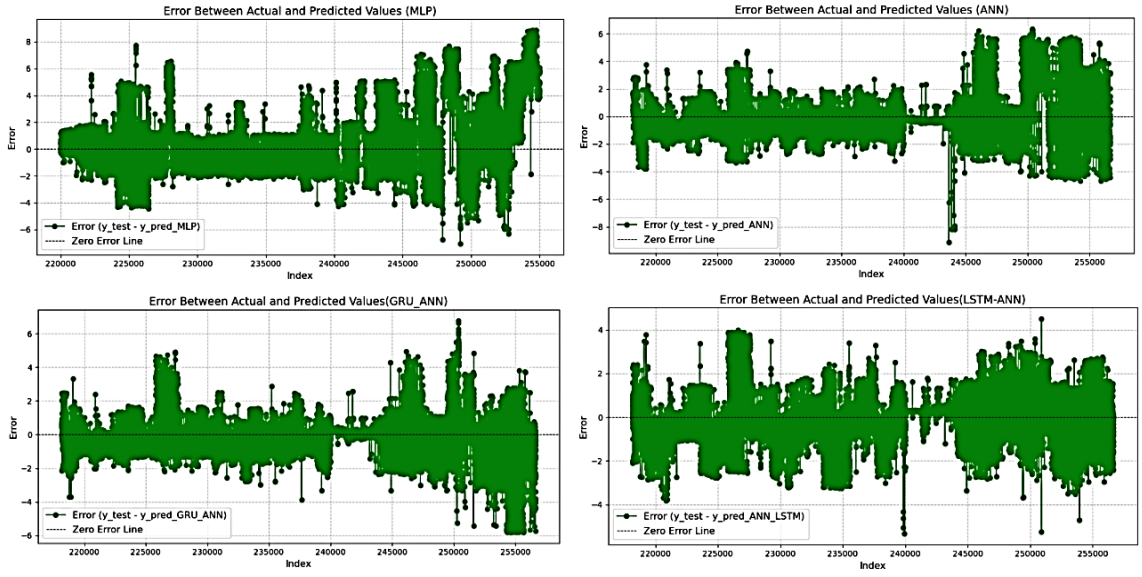


Fig. 5. Error analysis between actual and predicted values for different models.

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The "Smart Connected Mine" project, which has received support from the Moroccan Ministry of Higher Education, Scientific Research and Innovation, the Digital Development Agency (DDA), and the National Center for Scientific and Technical Research of Morocco (CNRST) through the Al- Khawarizmi program, is the context for this research. The work presented in this article was part of a collaborative effort among various partners, including MASCIR (Moroccan Foundation for Advanced Science, Innovation and Research), REMINEX Engineering, RD and Project Management subsidiary of MANAGEM group, ENSIAS (National School of Computer Science and Systems Analysis), UCA (Cadi Ayyad University), and ENSMR (National School of Mines of Rabat). We express our gratitude to MANAGEM Group and its subsidiary CMG for granting us permission to conduct on-site re- search and data collection as an industrial partner in this project.

References

[1] D. A. Narciso, F. Martins, Application of machine learning tools for energy efficiency in industry: A review, Energy Reports 6 (2020) 1181– 1199.

[2] F. S´anchez, P. Hartlieb, Innovation in the mining industry: Technological trends and a case study of the challenges of disruptive innovation, Mining, Metallurgy & Exploration 37 (5) (2020) 1385–1399.

[3] S. Qassimi, E. H. Abdelwahed, Disruptive innovation in mining industry 4.0, Distributed Sensing and Intelligent Systems: Proceedings of ICDSIS 2020 (2021) 313–325.

[4] Z. Hyder, K. Siau, F. Nah, Artificial intelligence, machine learning, and autonomous technologies in mining industry, Journal of Database Management (JDM) 30 (2) (2019) 67–79.

[5] A. Bendaouia, E. H. Abdelwahed, S. Qassimi, A. Boussetta, A. Benhayoun, I. Benzakour, O. Amar, Y. Zennayi, F. Bourzeix, K. Ba¨ına, et al., Digital transformation of the flotation monitoring towards an online analyzer, in: International Conference on Smart Applications and Data Analysis, Springer, 2022, pp. 325–338.

[6] S. Es-sahly, A. Elbasbas, K. Naji, B. Lakssir, H. Faqir, S. Dadi, R. Rabie, Nir-spectroscopy and machine learning models to pre-concentrate copper hosted within sedimentary rocks, Mining, Metallurgy & Exploration 41 (4) (2024) 1979–1995.

[7] M.-J. Li, W.-Q. Tao, Review of methodologies and polices for evaluation of energy efficiency in high energy-consuming industry, Applied Energy 187 (2017) 203–215.

[8] J. T. McCoy, L. Auret, Machine learning applications in minerals pro- cessing: A review, Minerals Engineering 132 (2019) 95–109

[9] O. Hasidi, E. H. Abdelwahed, A. Qazdar, A. Boulaamail, M. Krafi, I. Benzakour, F. Bourzeix, S. Ba¨ına, K. Ba¨ına, M. Cherkaoui, et al., Digital twins-based smart monitoring and optimisation of mineral processing industry, in: International Conference on Smart Applications and Data Analysis, Springer, 2022, pp. 411–424.

[10] S. Ma, Y. Zhang, J. Lv, Y. Ge, H. Yang, L. Li, Big data driven predictive production planning for energy-intensive manufacturing industries, Energy 211 (2020) 118320.

[11] B. Mahesh, Machine learning algorithms-a review, International Journal of Science and Research (IJSR), 9 (1) (2020) 381–386.

[12] D. Radchenko, A. Bondarenko, Mining engineering system as an energy asset in industry 4.0, in: E3S Web of Conferences, Vol. 58, EDP Sciences, 2018, p. 01009.

[13] A. Rihi, S. Ba¨ına, F.-z. Mhada, E. Elbachari, H. Tagemouati, M. Guer- boub, I. Benzakour, Predictive maintenance in mining industry: Grinding mill case study, Procedia Computer Science 207 (2022) 2483–2492.

[14] M. Imam, K. Ba¨ına, Y. Tabii, E. M. Ressami, Y. Adlaoui, I. Benzakour,E. H. Abdelwahed, The future of mine safety: a comprehensive review of anti-collision systems based on computer vision in underground mines, Sensors 23 (9) (2023) 4294.

[15] K. Clero, M. Nadour, S. Ed-Diny, M. Achalhi, M. Cherkaoui, H. Ait Ab- delali, S. El Fkihi, I. Benzakour, S. Rziki, H. Tagemouati, et al., A review of the use of thermal imaging and computer vision for pattern recognition, Computer Science & Information Technology (CS & IT) 13 (21) (2023).

[16] T. Igogo, K. Awuah-Offei, A. Newman, T. Lowder, J. Engel-Cox, Inte- grating renewable energy into mining operations: Opportunities, challenges, and enabling approaches, Applied Energy 300 (2021) 117375.

[17] J. Jeswiet, A. Szekeres, Energy consumption in mining comminution, Procedia CIRP 48 (2016) 140–145.

[18] A. Dasgupta, A. Nath, Classification of machine learning algorithms, International Journal of Innovative Research in Advanced Engineering (IJIRAE) 3 (3) (2016) 6–11.

[19] C. Loudari, M. Cherkaoui, R. Bennani, I. El Harraki, O. Fares, M. El Ad- nani, I. Benzakour, F. Bourzeix, K. Baina, et al., Predicting grinding mill power consumption in mining: A comparative study, in: 2023 7th IEEE Congress on Information Science and Technology (CiSt), IEEE, 2023, pp. 395–399.

[20] J. Peng, W. Sun, J. Xu, G. Zhou, L. Xie, H. Han, Y. Xiao, J. Chen, Q. Li, Analyzing process parameters for industrial grinding circuit based on machine learning method, Advanced Powder Technology 34 (9) (2023)104113.

[21] S. Avalos, W. Kracht, J. M. Ortiz, Machine learning and deep learning methods in mining operations: A data-driven sag mill energy consumption prediction application, Mining, Metallurgy & Exploration 37 (2020) 1197–1212.

[22] A. Tohry, S. C. Chelgani, S. Matin, M. Noormohammadi, Power-draw prediction by random forest based on operating parameters for an industrial ball mill, Advanced Powder Technology 31 (3) (2020) 967–972.

[23] G. Liu, K. Wang, X. Hao, Z. Zhang, Y. Zhao, Q. Xu, Sa-lstms: A new advance prediction method of energy consumption in cement raw materials grinding system, Energy 241 (2022) 122768.

[24] J. Liu, Q. Zhang, Z. Dong, X. Li, G. Li, Y. Xie, K. Li, Quantitative evaluation of the building energy performance based on short-term energy predictions, Energy 223 (2021) 120065.

[25] O. Bascur, A. Soudek, Grinding and flotation optimization using op- erational intelligence, Mining, Metallurgy & Exploration 36 (1) (2019)139–149.

[26] R. Rajamani, P. Kumar, N. Govender, The evolution of grinding mill power models, Mining, Metallurgy & Exploration 36 (1) (2019) 151–157.

[27] Z. Ghasemi, F. Neumann, M. Zanin, J. Karageorgos, L. Chen, A com- parative study of prediction methods for semi-autogenous grinding mill throughput, Minerals Engineering 205 (2024) 108458.

[28] G. Lin, A. Lin, J. Cao, Multidimensional knn algorithm based on eemd and complexity measures in financial time series forecasting, Expert Systems with Applications 168 (2021) 114443.

[29] P. C. d. L. e Silva, C. A. S. Junior, M. A. Alves, R. Silva, M. W. Cohen,F. G. Guimar˜aes, Forecasting in non-stationary environments with fuzzy time series, Applied Soft Computing 97 (2020) 106825.

[30] R. Gao, L. Du, K. F. Yuen, Robust empirical wavelet fuzzy cognitive map for time series forecasting, Engineering Applications of Artificial Intelligence 96 (2020) 103978. 177(4325-4343).

[19] Varfinezhad, R., & Oskooi B. (2020). 2D DC resistivity forward modeling based on the integral equation method and a comparison with the RES2DMOD results. Journal of Earth and Space Physics.

[20] Ekinci, Y. L., Demirci, A., & Ertekin, C. (2008). Delineation of the seawater-freshwater interface from the coastal alluvium of Kalekoy-Gokceada, NW Turkey. *Journal of Applied Sciences*, 8, 1977-1981.

[21] Ekinci, Y. L., & Demirci, A. (2008). A damped least-squares inversion program for the interpretation of Schlumberger sounding curves. Journal of Applied Sciences, 8, 4070–4078.

[22] Temel, R.O., & Ciftci, N.B. (2002). Stratigraphy and depositional environments of the tertiary sedimentary units in Gelibolu Peninsula and islands of Gokceada and Bozcaada (Northern Aegean Region, Turkey). TAPG Bulletin, 14(2), 17-40.