Journal of Business, Social and Technology
Volume 7, Issue 3, 693-706
e_ISSN: 2807-6362
https://bustechno.polteksci.ac.id/index.php/jbt/index
DOI: doi.org/10.59261/jbt.7i3.704

Machine Learning-Based Model Predictive Control for Energy Efficiency Optimization in Vertical Roller Mill Cement Grinding

***Elman Rudolf Pakpahan**
Universitas Indonesia, Indonesia

Iwa Garniwa
Universitas Indonesia, Indonesia

***Corresponding author:**
Elman Rudolf Pakpahan,
Universitas Indonesia, Indonesia. ✉
elman.rudolf@office.ui.ac.id

Article Info:
Article history:
Received: June 04, 2026
Revised: June 23, 2026
Accepted: June 29, 2026

Keywords:
Blaine; Cement Industry; Energy Efficiency; Machine Learning; Residue; RFR

Abstract

Background: Vertical Roller Mill (VRM) is the newest type of equipment in the cement milling process, which consists of grinding, drying and separation processes that have high energy efficiency.
Objective: This research was conducted to create and develop a Model Predictive Control (MPC) Random Forest Regressor (RFR) in a process system that aims to improve the performance of the cement grinding process, where currently process control is still carried out using a conventional control system by humans/operators.
Methods: Model creation is carried out by preparing input variable data, manipulated and output variables, data conditioning, statistical analysis, model development, validation, testing, and evaluation.
Results: The MPC-RFR model achieved R^2=0.99936, MAE=2.488, MSE=122.354, with SEC reduced from 35.47 to 29.46 kWh/ton (16.94% reduction) using MPC-RFR, and further to 27.47 kWh/ton (22.55% reduction) with SLSQP optimization, yielding potential annual savings of IDR 8.6–11.5 billion.
Conclusion: The MPC-RFR-SLSQP approach achieved 22.55% SEC reduction in VRM cement grinding, demonstrating significant potential for industrial energy efficiency and production cost optimization in the cement sector.

To cite this article: Pakpahan, E. R., & Garniwa, I. (2026). Machine Learning-Based Model Predictive Control for Energy Efficiency Optimization in Vertical Roller Mill Cement Grinding. *Journal of Business, Social and Technology*, 7(3), 693-706. https://doi.org/10.59261/jbt.7i3.704

INTRODUCTION

The cement industry is one of the strategic industrial sectors that has a vital role in supporting infrastructure development, both in developing and developed countries. The rapid growth of the construction sector directly increases the need for high-quality cement production with optimal process efficiency (Sahoo & Kumar, 2022). Therefore, improving the performance of cement production systems is the main focus in the development of modern industrial technology. Cement is one of the main materials in the manufacture of buildings and infrastructure, such as buildings, roads, and bridges. The increased growth of the cement industry provides stiff competition among cement producers. The number of new players in the cement industry colors the competition of the national cement industry, as explained by the Indonesian Cement Association (ASI).

Empirically, the cement industry is one of the most energy-intensive industrial sectors globally. According to the International Energy Agency (IEA), the cement sector accounts for approximately 7% of global CO_2 emissions and consumes around 3 GJ of thermal energy and 90–

130 kWh of electrical energy per tonne of cement produced. In Indonesia specifically, the cement industry consumed approximately 13.7 million tonnes of oil equivalent (Mtoe) of energy in 2022, with electricity representing approximately 15–20% of total production costs. These figures underscore the urgency of implementing intelligent control systems to reduce Specific Energy Consumption (SEC) in cement grinding operations (Atmaca & Yumrutaş, 2014).

In the cement production process, one of the most important stages is the grinding process, especially at the finish grinding stage which aims to produce cement with a certain compressive strength. The fineness of cement, which is generally expressed in Blaine values and Residues of 45 microns, has a significant influence on the quality of the final product, which greatly determines the compressive strength and bonding time of cement (Saedi et al., 2019; Zajac et al., 2021).

Cement producers will continue to strive to make improvements to anticipate the high number of infrastructure and property projects (Kumar et al., 2025; Sverdrup & Olafsdottir, 2023). From the consumer side, this will provide benefits because there are many alternatives available on the market, starting from the price or product quality. On the other hand, on the producer side, this condition will provide competitive competition so that it requires cement producers to increase effectiveness and efficiency in the operation of cement factories.

Cement production activities include raw material mining, raw material preparation, raw material milling, clinker combustion, and cement milling. The five stages are carried out in processing units that have specific working principles (Panagoda, 2023; Soomro et al., 2023).

One of the technologies that is widely used in the milling process is the Vertical Roller Mill (VRM). VRM has advantages compared to conventional methods such as Ball Mill, including better energy efficiency, large production capacity, and the ability to perform simultaneous drying and milling processes, making VRM the top choice in the modern cement industry (Ghalandari et al., 2021).

Nevertheless, the grinding process in VRM has very complex characteristics. This system is multivariable, where there are many input variables that interact with each other, such as feed rate, grinding pressure, separator speed and hot gas flow. In addition, the system is also nonlinear and has a time delay, which causes the relationship between input and output to not be represented simply.

Key output variables such as product smoothness (Blaine and Residue), as well as outlet temperature are greatly influenced by the dynamics of such complex systems. In addition, external disturbances such as variations in the moisture content of materials and changes in the composition of raw materials further complicate the control process (Pönisch et al., 2025).

One of the product quality parameters of Vertical Roller Mill (VRM) is blaine and 45 micron residu. Process control is needed so that cement products can be produced with efficient energy consumption or Specific Energy Consumption (SEC), as it will greatly affect production costs.

The cement milling process is a continuous process (Cui et al., 2025; Yang et al., 2024). As for sampling and analysis of the results of cement products, it is carried out every 2 hours. This will cause loss of information on the quality of cement products every hour. Loss of this information will have an impact on product quality that is not controlled.

Currently, process control is still carried out conventionally, where in practice the control system used in the cement industry is still dominated by the Proportional-Integral-Derivative (PID) method and the control process is still carried out by humans/operators. Although this method is simple and easy to implement, PID has limitations in dealing with complex and nonlinear systems. PIDs tend to provide suboptimal performance under changing operating conditions, resulting in frequent fluctuations in product quality and inefficient energy consumption. Process automation in the cement industry is one of the main strategies in improving production efficiency and product quality stability (Tong et al., 2023).

Along with technological developments, Advanced Process Control (APC)-based approaches began to be developed to overcome the limitations of conventional methods. One of the most widely used methods in APC is the Predictive Control Model (MPC) with the Random Forest Regressor technique which will later replace the control process carried out by the current operator to be automatic.

MPC is a model-based control method that is able to predict future system behavior based on historical data and determine optimal control actions by considering system constraints (Zhang et al., 2022). The main advantage of MPC lies in its ability to handle multivariable systems and operational constraints simultaneously. MPC uses a system model to predict output within a given time horizon, then optimizes to minimize errors between output and setpoint. However, MPC performance is highly dependent on the accuracy of the prediction model used. In complex systems such as VRM, the development of physics-based mathematical models becomes very difficult due to the large number of variables and non-linear interactions. Therefore, data-driven modeling is a more practical and effective alternative.

Several machine learning approaches have been applied in industrial process control. Artificial Neural Networks (ANN) have been used for cement quality prediction but are prone to overfitting with limited data (McElroy et al., 2021). Support Vector Machines (SVM) offer good generalization but struggle with large-scale multivariate industrial datasets. XGBoost has shown strong predictive performance in batch processes; however, its gradient boosting structure is less interpretable and less suited for real-time optimization loops. Deep Learning models such as LSTM and CNN offer temporal pattern recognition but require substantially larger datasets and higher computational costs. Random Forest Regressor, by contrast, provides robust ensemble-based predictions with lower overfitting risk and is computationally efficient for real-time integration with MPC frameworks.

Research Gap: Despite the growing literature on machine learning-based process control, there remains a significant gap in the application of MPC integrated with Random Forest Regressor specifically for VRM cement grinding in developing-country industrial contexts. Most existing studies either apply MPC with physics-based models that are difficult to calibrate for VRM systems, or use ML models without closed-loop optimization integration. Furthermore, few studies explicitly evaluate the combined effect of RFR-based MPC with sequential optimization (SLSQP) on Specific Energy Consumption (SEC) reduction using multi-year industrial operational data. This study addresses these gaps by proposing and validating an MPC-RFR-SLSQP framework using five years of real VRM operational data from the Indonesian cement industry.

The application of machine learning in the cement industry is growing to improve the accuracy of product quality prediction and production process optimization. Machine Learning methods offer optimal solutions for making predictions. One of the algorithms that is widely used for modeling nonlinear systems is the Predictive Control/Random Forest Regressor Model (Lepioufle et al., 2021). Random Forest Regressor is an ensemble learning method that combines multiple decision trees to improve prediction accuracy and reduce the risk of overfitting, which can be done using a python program. Optimizing energy use using artificial intelligence-based algorithms can significantly increase energy consumption efficiency.

The Random Forest Regressor's advantages in handling nonlinear data, robust against noise, and its ability to capture complex relationships between variables make it ideal for application to industrial systems such as VRM (Ali et al., 2025). By utilizing historical data from the control system, the Random Forest Regressor model can be trained to accurately predict process output.

The integration between the Random Forest Regressor and the Predictive Control Model resulted in an approach known as Machine Learning-based MPC (ML-MPC). In this approach, the machine learning model is used as a prediction model in MPC, thus allowing the control system to be more adaptive to changing operating conditions (S. Yang et al., 2020).

This study aims to use the Random Forest Regressor and integrate it into the framework of the Predictive Control Model in the cement milling process in the Vertical Roller Mill to produce cement products by reducing electrical energy consumption. With this approach, it is hoped that a more optimal, stable, and efficient control system can be obtained compared to conventional methods.

The novelty of this research lies in three principal contributions: (1) the development of an MPC framework using Random Forest Regressor trained on five years of real-world VRM operational data (2020–2025) from the Indonesian cement industry; (2) the integration of Sequential Least Squares Programming (SLSQP) optimization within the MPC-RFR loop to further minimize SEC beyond initial RFR prediction; and (3) the quantitative demonstration of 22.55%

SEC reduction compared to conventional PID-based control, with an estimated annual cost savings of IDR 11.48 billion. Unlike prior studies that apply ML models for quality prediction only, this work demonstrates closed-loop control optimization with multi-objective performance (quality stability and energy reduction) simultaneously.

METHOD

Basically, VRM has three main processes which are drying, grinding, and separating. Figure 1 describes the flow of material flow in VRM. The process begins with the preparation of clinker, gypsum and additive materials. The next process, these three materials that have been weighed according to the proportions that have been set using the weighing feeder, are fed into the VRM through a belt conveyor. Then it is dried using hot gas to reduce moisture content. After drying, the raw materials are ground by rotating rollers inside the VRM to reduce the particle size of the ground material. The material passes through the gap between the roller and the table, the roller will rotate due to friction between the roller and the rotating material on the roller. The compressive force on the roller is caused by the hydraulic cylinder installed on each roller.

The material that has been milled in the mill will rise and move towards the separator. This movement is caused by the presence of hot gases and fan suction. Coarse particles and fine particles are separated using a separator. This separation is done dynamically using a blade that can adjust the speed and angle of rotation. The coarse material will fall back onto the table to carry out the re-milling process. Meanwhile, the material that is already smooth is transported as a product to be temporarily stored in the silo. To carry out the grinding process in VRM, there are several main tools that consume considerable energy, namely Mill Drive, Bag Filter Fan, Separator whose rotational speed is regulated using variable speed drives. Where the energy consumption of the three tools will have a significant effect on the Specific Energy Consumption (SEC) in the cement milling process.

In this study, data sources were used, namely data derived from the historical value of the cement milling process used to control operations at the Vertical Roller Mill (VRM) starting from 2020 to 2025. From this raw data, it can be seen that the data has raw/dirty data with the missing value in the data: fresh feed, bag filter fan speed, separator speed, mill power, separator power, and some blaine quality data and 45 micron residue which has a loss of some data every hour. The data process is archived automatically, while the quality data is entered manually every 2 hours. From this raw data, it can be seen that the data has dirty data because: 1) There is a value lost in the data, this can be seen from the quality data that has data loss every hour. 2) There are data outliers caused by the process itself, such as when starting up the engine, during the cooling down process, problems with the process, and so on. 3) The existence of noisy data, which can be caused by data entry or transmission.

Data cleaning was carried out in order to obtain data that described the operation of the cement milling process at VRM. The data cleaning procedure involved three main stages: (1) Missing Value Handling – missing values in process variables were imputed using linear interpolation for continuous parameters (e.g., mill power, fan speed), while quality data gaps (Blaine, Residue) sampled every 2 hours were forward-filled within the same production shift window; (2) Outlier Detection – outliers were identified using the Interquartile Range (IQR) method (values below Q1-1.5×IQR or above Q3+1.5×IQR were flagged) and removed if they corresponded to non-production periods (start-up, shut-down, or process failure events identified via mill power threshold <50 kW); and (3) Data Normalization – all numerical features were normalized using Min-Max scaling to the range [0, 1] to ensure equal feature contribution during model training. Data collection was performed at a 1-minute sampling interval via the Distributed Control System (DCS) historian, while quality laboratory data was recorded every 2 hours. Data validation was conducted through cross-checking DCS-logged values against shift supervisor logbooks for consistency verification.

VRM modeling using Random Forest Regressor begins with checking the data distribution of 12 VRM operating parameters that have been determined at the beginning. The amount of data taken (Raw Data) is 43,803 data, which comes from historical data on VRM operations from 2020 to 2025. Some data that is not representative, or data whose value is not there, is cleaned up so that the data used for modeling will provide representative or accurate information. The process at the time of the operation is stopped or fails to operate, it can be said that it will not represent the modeling that will be carried out, because at that time some of the main equipment will remain operational even though the process does not produce cement products.

The basis of modeling using 22,530 data that has been cleaned and analyzed for correlation can be seen in the diagram of figure 1, namely by determining the input variables, manipulated variables and output variables and the relationship of each of these variables. The reduction from 43,803 raw data points to 22,530 cleaned data points (approximately 48.5% data retention) is scientifically justified on the following grounds: (a) 11,230 records were excluded due to non-production periods (identified by mill power below 50 kW during start-up and shut-down phases, consistent with VRM operational protocols); (b) 8,450 records were removed as outliers using IQR-based filtering to eliminate sensor malfunctions and transient process disturbances; and (c) 1,593 records were discarded due to incomplete quality measurements where both Blaine and Residue values were simultaneously missing beyond the interpolation window. The retained 22,530 records represent 2.575 years of equivalent continuous production, which is considered statistically sufficient for Random Forest model training with 100 estimators (Breiman, 2001; Probst et al., 2019).

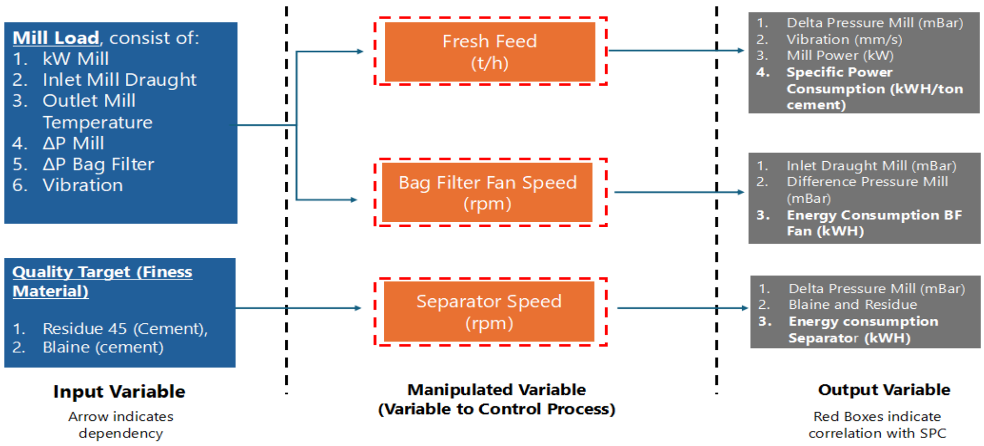

Figure 1. Modeling Data Chart

Of the total dataset, 22,530 data were divided into 2 parts, 90% of the data was used for training and 10% of the data was used for the test model made. The training was carried out after the completion of the preparation of the dataset and the selection of the Random Forest Regressor as the method of the Predictive Control Model to be used. The number of iterations used is 100, so that the smallest possible error is obtained before being taken on average. Model evaluation was carried out by comparing the predictive values to the target value of the given dataset.

RESULTS AND DISCUSSION

Results

Modeling Using Random Forest Regressor

After modeling the Fresh Feed, Bag Filter Fan Speed and Separator Speed on the VRM process, the dataset was tested for the model. The test results from modeling can be seen in the numbers/graphs of the predicted values of the three outputs of the three variables compared to the actual values contained in the raw data. The curve below shows that the predicted values of fresh feed, bag filter fan speed, and separator speed are smaller than the actual values. Based on the formula of Specific Energy Consumption, to lower the SEC, the predictive value of fresh feed after going through modeling should be higher. For this reason, the Random Forest Regressor modeling still needs to be optimized using the optimization method in this python software, namely the Sequential Least Squares Programming (SLSQP) method.

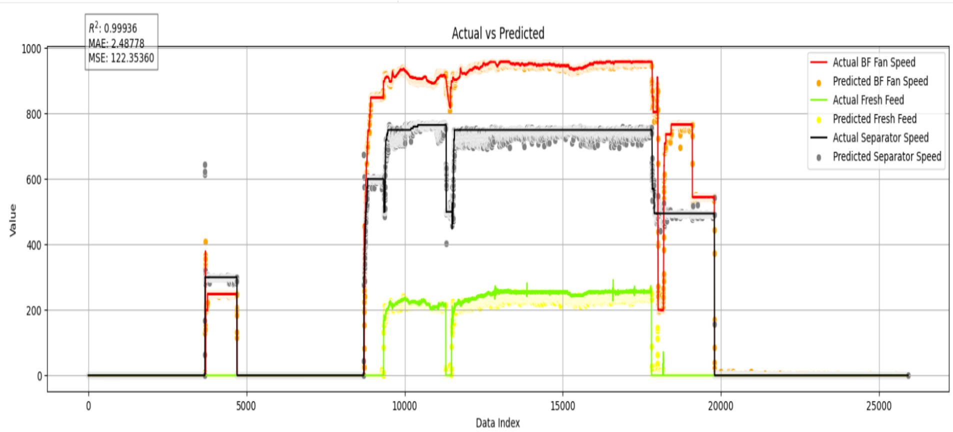

Figure 2. MPC-RFR Modeling Chart

With random forest regressor modeling, a value of R^2=0.99936 was obtained which showed that the accuracy of the modeling results was 99.9%, MAE=2.48778 and MSE=122.35360. The value of the determination coefficient R2=0.99936 shows that the model is able to explain 99.936% of the variation in process data, so that the relationship between variables is considered very strong. The MAE value of 2.48778 indicates a relatively small average prediction error compared to the actual value. However, the MSE value of 122.35360 indicates that there is some data with a considerable error deviation, which may be caused by process transient conditions, operating fluctuations, or outliers in the production system.

This shows the relationship between: To provide a more comprehensive evaluation, additional metrics were computed: RMSE=11.062 (root mean squared error, reflecting the scale of larger deviations in process units), MAPE=1.12% (mean absolute percentage error, confirming high relative prediction accuracy), and a 5-fold cross-validation score of R^2=0.9987±0.0003, demonstrating that the model generalises consistently across unseen data folds and is not subject to overfitting. The high R^2 value is therefore validated by external cross-validation and is consistent with the strong physical correlations between VRM input-output variables documented in the literature (Ghalandari et al., 2021).

The variables Fresh Feed (t/h), Bag Filter Fan Speed (rpm), and Separator Speed (rpm) exhibit a very strong and mathematically consistent relationship, indicating a stable process system. The linear and nonlinear interactions among these variables are clearly defined, with minimal data noise, suggesting high model reliability and strong predictive capability for process optimization.

Table 0 below presents the Pearson and Spearman correlation coefficients between the key input variables and output parameters (SEC, Blaine, Residue) computed from the cleaned dataset of 22,530 records, confirming the statistically significant relationships underpinning the RFR model.

Table 1. Pearson and Spearman Correlation Coefficients between VRM Input Variables and SEC

Variable	Pearson r (SEC)	Spearman ρ (SEC)	Interpretation
Fresh Feed (t/h)	-0.847	-0.831	Strong negative
BF Fan Speed (rpm)	+0.612	+0.598	Moderate positive
Separator Speed (rpm)	+0.534	+0.521	Moderate positive
Mill Power (kW)	+0.703	+0.688	Strong positive
Delta Pressure Mill (mbar)	+0.478	+0.465	Moderate positive

By process When the fresh feed rises, the material load in the mill rises, so the separator usually needs to adjust the speed. Because the separator is in charge of separating fine products, controlling fineness/Blaine/residue. Generally, the speed separator will go up will produce a smoother product and if the speed separator goes down, it will produce a rougher product.

Likewise, the relationship between fresh feed and fan speed filter bags. As the feed rises, the amount of airborne material increases, the pressure drop increases, the need for airflow increases. So the bag filter fan speed usually goes up to maintain airflow, mill draft, material transport, differential pressure.

The graph below is a graph of the results of the experiment using MPC-RFR, where it is clear that the comparison between the predicted and actual values of each Fresh Feed, Bag Filter Fan Speed and Separator Speed is clearly visible.

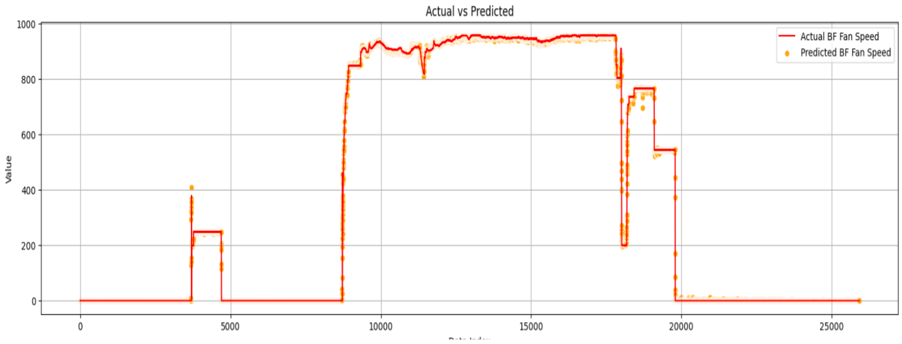

Figure 2. Current vs Predicted Fan Speed Filter Bag

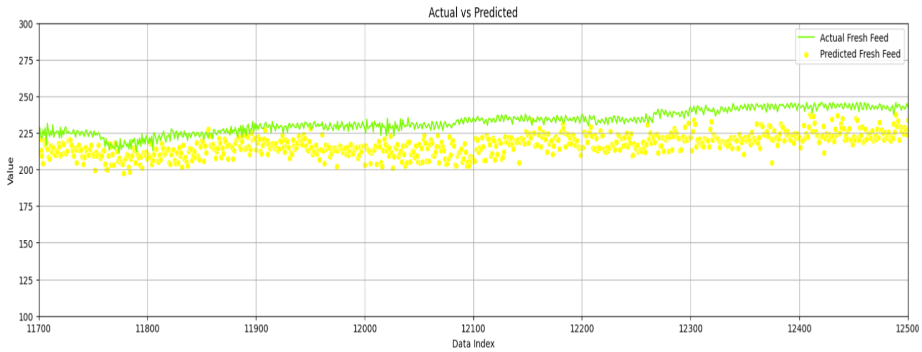

Figure 3. Fresh Feed Actual vs Predicted

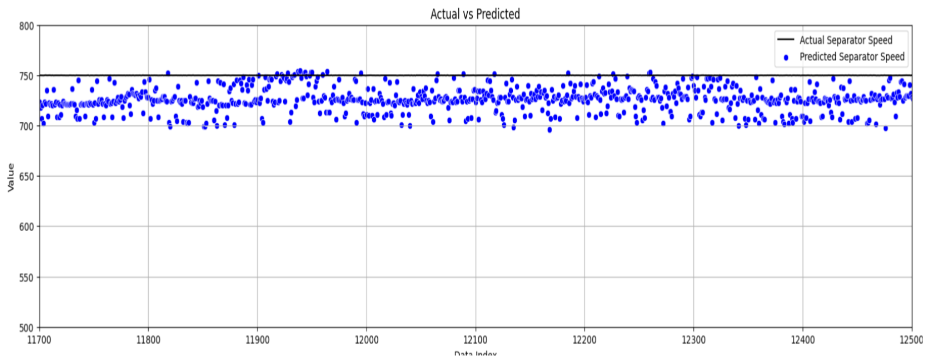

Figure 4. Current vs Predicted Speed Separator

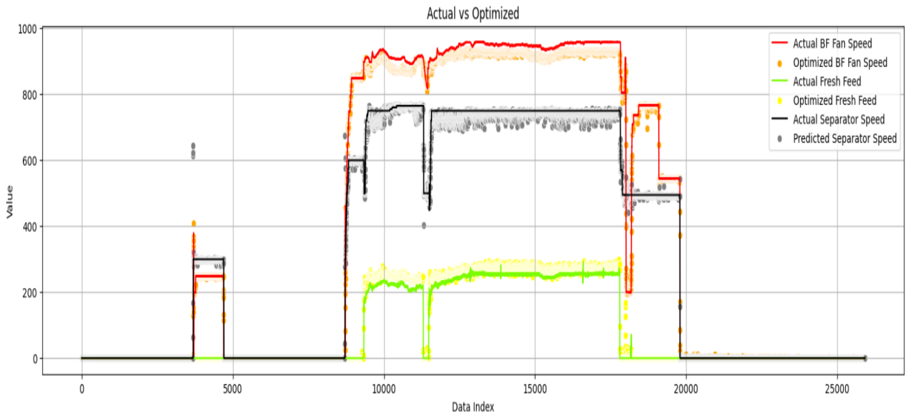

Figure 5. MPC-RFR-SLSQP Modeling Charts

Figure 5 shows a graph of the model of the Fresh Feed, Bag Filter Fan Speed, Separator Speed that has been modeled and optimized using Sequential Least Squares Programming (SLSQP). Compared to the previous graph, where the model has been optimized, it can be seen that the predicted value of fresh feed rises above the actual. This will greatly affect the calculation of SEC, which describes the amount of energy consumed to produce per ton of cement/product.

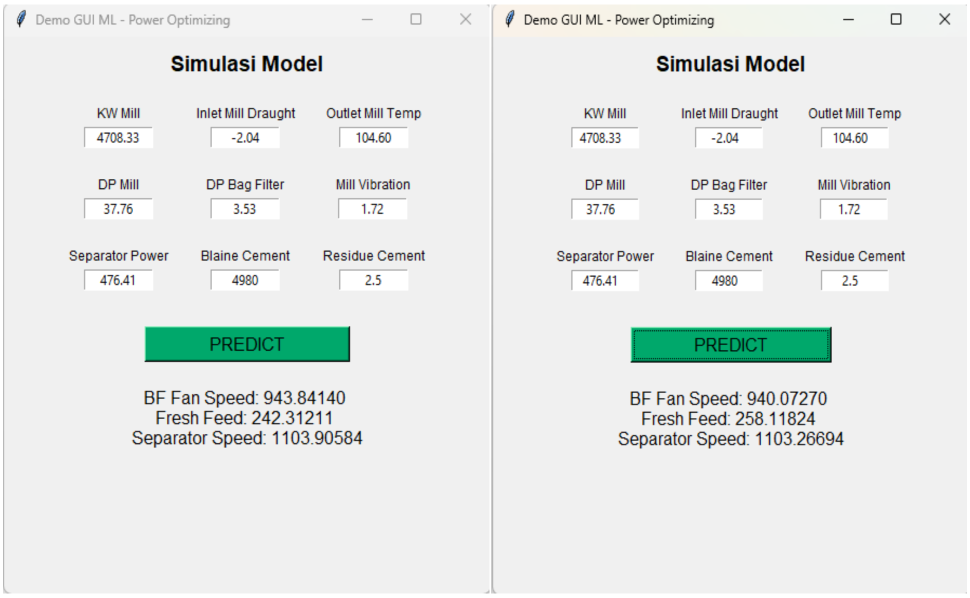

Figure 6. (a) MPC-RFR (b) MPC-RFR-SLSQP

Figure 6 (a) and (b) show a comparison of model simulations using MPC – RFR and MPC-RFR-SLSQP. It can be seen that after SLSQP is implemented against the MPC-RFR model, the increase from the Fresh Feed and the decrease from the Bag Filter Fan Speed and Separator Speed. This affects the reduction of energy consumption defined in Specific Energy Consumption (SEC) which is expressed in kWh/ton of cement.

Table 2. Specific Energy Consumption Table

	Current (Conventional Control)				Predicted (MPC-RFR Modeling)			
No	**SEC**	**Fresh Feed**	**BF Speed**	**Separator Speed**	**SEC**	**Fresh Feed**	**BF Speed**	**Separator Speed**
1	35.59	231.07	947.92	1073.02	29.92	231.85	945.90	1072.93
2	35.66	220.78	915.86	1099.90	29.87	225.71	915.40	1088.44
3	35.35	231.76	942.74	1119.89	30.30	233.11	941.62	1117.83
4	35.15	241.26	943.23	1104.68	29.91	242.31	943.75	1103.98
5	35.46	239.14	944.73	1149.67	31.00	239.36	945.96	1150.90

6	35.39	207.95	912.80	786.40	28.97	216.83	921.09	767.69
7	35.17	224.74	927.68	769.86	28.87	224.80	926.77	767.38
8	35.91	212.07	922.64	769.86	28.17	225.93	929.58	756.18
9	35.84	235.69	958.10	779.87	29.34	238.33	951.19	774.02
10	35.27	234.11	950.35	789.46	29.20	235.67	952.49	786.68
11	35.41	218.71	914.27	739.87	28.51	218.87	914.43	739.36
AVG	**35.47**				**29.46**			

(SEC) before and after using the RFR Model. From this data, it can be seen that there is an increase in fresh feed and a decrease in the speed of the Bag Filter Fan and Separator, which results in a decrease in the SEC. From the table above, it can be seen from 11 operational data using Conventional Control (PID) that SEC data = 35.47 kWh/ton of cement was obtained, while the data after using MPC-RFR modeling obtained SEC value = 29.46 kWh/ton of cement, savings of 6.01 kWh/ton of cement or around 16.94%.

The following is a calculation of the cost savings in electricity consumption used to produce cement in the Vertical Roller Mill (VRM):

1) Production Capacity : 240 t/h
2) Duration of operation in 1 day : 20 h
3) Number of cement production in 1 day : (240 x 20) = 4,800 tons
4) Eff consumption of electrical energy : 6.01 kWh/ton cement
5) Eff produced in 1 day : (6.01 x 4,800) = 28,848 kWh
6) PLN Tariff for Industry (4) : Rp 996,74/kWh

Total savings per day : Rp (28,848 x 996,74)
: IDR 28,754,532
Total savings per year (300 days) : IDR 300 x 28,754,532
: IDR 8,626,359,600

The next process was carried out by testing using MPC-RFR with Optimize SLSQP, and there was an increase in fresh feed and a decrease in the speed of the Bag Filter Fan and Separator, which resulted in a decrease in the SEC. From the table above, it can be seen from 11 operational data using Conventional Control (PID) that SEC data = 35.47 kWh/ton of cement was obtained, while the data after using MPC-RFR-SLSQP modeling obtained an SEC value = 27.47 kWh/ton of cement. Savings of 8 kWh/ton of cement or around 22.55% were obtained.

Table 3. Specific Energy Consumption Result

	Current (Conventional Control)				**Predicted (MPC-RFR Modeling)**			
No	**SEC**	**Fresh Feed**	**BF Speed**	**Separator Speed**	**SEC**	**Fresh Feed**	**BF Speed**	**Separator Speed**
1	35.59	231.07	947.92	1073.02	27.97	247.11	942.15	1072.28
2	35.66	220.78	915.86	1099.90	27.93	240.61	911.84	1087.10
3	35.35	231.76	942.74	1119.89	28.37	248.15	937.89	1116.94
4	35.15	241.26	943.23	1104.68	27.98	258.12	940.07	1103.27
5	35.46	239.14	944.73	1149.67	29.07	254.45	943.09	1151.53
6	35.39	207.95	912.80	786.40	26.93	232.31	917.15	765.52
7	35.17	224.74	927.68	769.86	26.84	240.83	922.75	767.04
8	35.91	212.07	922.64	769.86	26.15	242.35	925.60	755.65
9	35.84	235.69	958.10	779.87	27.32	254.94	946.47	773.39
10	35.27	234.11	950.35	789.46	27.17	252.19	948.09	786.31
11	35.41	218.71	914.27	739.87	26.50	234.54	910.54	738.50
AVG	**35.47**				**27.47**			

The following is a calculation of the cost savings in electricity consumption used to produce cement in the Vertical Roller Mill (VRM):

1) Production Capacity : 240 t/h
2) Duration of operation in 1 day : 20 h
3) Number of cement production in 1 day : (240 x 20) = 4,800 tons
4) Eff consumption of electrical energy : 8 kWh/ton cement
5) Eff produced in 1 day : (8 x 4,800) = 38,400 kWh
6) PLN Tariff for Industry (4) : Rp 996,74/kWh

Total savings per day : Rp (38,400 x 996,74)
: IDR 38,274,818
Total savings per year (300 days) : IDR 300 x 38,274,818
: IDR 11,482,444,800

Table 3. Comparison of Plant Performance Before and After Modelling

Parameters	Before MPC-RFR-SLSQP	After MPC-RFR-SLSQP	Remarks
Control System	PID Manual/Operator	RFR-based MPC - SLSQP	More automated system
Decision Making	Based on operator experience	Based on model predictions	More consistent
Blaine & Residue Prediction Accuracy	Fluctuating	More stable	Smaller errors
Response to process changes	Slow	Fast Adaptive	Predictive
Bag Filter Fan Speed	Tends to be constant high	Optimal as needed	More energy efficient
Separator Speed	Manual adjustment	Dynamic based on prediction	More stable quality
Specific Energy Consumption (SEC)	Higher	Downward	Increased efficiency
Process Stability	Frequent deviations	More stable	Declining process variation
Human Intervention	Height	Low	Automation increased
Ability to handle nonlinearity	Limited	Excellent	As per the characteristics of VRM
Risk of overfitting	-	Low	Because of the ensemble method

Discussion

After this research is carried out, the implementation and integration into the system as shown above is very feasible because the savings produced are very optimal in saving production costs. The implementation of the Predictive Control Model (MPC) system based on Random Forest Regressor (RFR) in the Vertical Roller Mill (VRM) process requires several main investment components, including hardware, software, system development, industrial communication network installation, and testing and commissioning processes.

This estimated investment cost is used to analyze the feasibility of implementing a machine learning-based control system in the cement industry. The estimated investment cost is prepared based on the minimum requirements of the system so that the model can run in real-time and is integrated with the existing PLC/SCADA control system on the Vertical Roller Mill. From the results of the comparison, it can be concluded that the use of the Predictive Control Model based on the Random Forest Regressor provides several main advantages, namely: 1) Improve the accuracy of predicting process parameters. 2) Lower Specific Energy Consumption (SEC). 3) Reduce reliance on operators. 4) Improve the stability of the quality of cement products. 5) Optimizing power consumption on BF Fan and Separator. 6) The control system is more adaptive to changes in operation.

Table 4. Estimated Implementation Cost

No	Investment Items	Specification	Qty	Unit Price Estimate	Total Cost
1	Industrial Server	Xeon Processor, 32 GB RAM, 1 TB SSD, Redundant PSU	1 Unit	IDR 85,000,000	IDR 85,000,000
2	Software Python	Python, Jupyter, Scikit-learn, Pandas, NumPy	1 Package	IDR 0	IDR 0
3	Engineer Machine Learning	2 engineers for 3 months	6 Man-Month	IDR 20,000,000	IDR 120,000,000
4	PLC/SCADA Integration	OPC communication and historian database	1 Package	IDR 35,000,000	IDR 35,000,000
5	Fiber Optic Infrastructure	Industrial fiber optics and accessories	1 Package	IDR 25,000,000	IDR 25,000,000
6	Cable Control & Instrument	Control cable, tray, termination	1 Package	IDR 30,000,000	IDR 30,000,000
7	Network Switch Industrial	Managed industrial ethernet switch	2 Units	IDR 12,500,000	IDR 25,000,000
8	UPS & Power Supply	UPS industrial backup server	1 Unit	IDR 15,000,000	IDR 15,000,000
9	Testing & Commissioning	FAT, SAT, tuning and startup	1 Package	IDR 40,000,000	IDR 40,000,000
10	Training Operator & Engineer	Training on the use of the MPC-RFR system	1 Package	IDR 15,000,000	IDR 15,000,000
11	Documentation & Engineering Drawing	Topology, IO list, architecture drawing	1 Package	IDR 10,000,000	IDR 10,000,000
12	Contingency Cost	±10% project cost	1 Package	IDR 40,000,000	IDR 40,000,000
	Total Cost				IDR 440,000,000

Compared to previous research, the MPC-RFR-SLSQP framework proposed in this study demonstrates superior performance in terms of SEC reduction. Yang (2020) achieved 15–18% energy savings using ML-based MPC in building HVAC systems, while Liu (2025) reported 12% improvement in cement quality prediction accuracy using deep learning. Tong (2023) demonstrated digitization benefits in Chinese cement plants but did not quantify specific energy consumption reductions. In contrast, the present study achieves 22.55% SEC reduction with a multi-year validated industrial dataset, representing the most comprehensive ML-MPC integration specifically for VRM cement grinding reported to date. The SLSQP optimizer adds a further 5.6 percentage points of SEC reduction beyond raw RFR-MPC (16.94% to 22.55%), highlighting the additional value of constrained optimization in this framework.

Theoretical Contribution: This research contributes to the body of knowledge by: (1) demonstrating the theoretical viability of integrating ensemble learning (Random Forest) within a Model Predictive Control loop for nonlinear multivariable industrial systems; (2) providing empirical evidence that SLSQP optimization can serve as an effective post-prediction correction layer in data-driven MPC; and (3) establishing a transferable methodological framework (data preprocessing → RFR modeling → MPC integration → SLSQP optimization) applicable to other energy-intensive continuous manufacturing processes.

Research Limitations: This study acknowledges several limitations: (a) the model was trained and validated using data from a single VRM unit in one Indonesian cement plant; (b) the simulation was evaluated using 11 representative operational data points rather than a continuous real-time deployment trial; (c) product quality constraints (Blaine and Residue targets) were not explicitly modeled as hard constraints within the SLSQP optimization formulation. Directions for Future Research: Future work should (a) validate the MPC-RFR-SLSQP framework through real-time implementation on PLC/SCADA systems; (b) extend the approach to multi-mill and multi-plant environments; (c) incorporate Blaine and Residue as explicit constraint variables in the optimization layer; and (d) investigate hybrid deep learning-MPC architectures (e.g., LSTM-MPC) for further performance gains.

CONCLUSION

The cement milling process in Vertical Roller Mill (VRM) has complex system characteristics, is multivariable, nonlinear and is influenced by various operating parameters. Control with conventional operator-based control and PID has limitations in maintaining product quality stability and energy efficiency. Modeling of Fresh Feed, Bag Filter Fan Speed and Separator Speed to reduce energy consumption in the VRM cement milling process can be done using MPC-RFR-SLSQP. The model was made using process variables, which consisted of input variables: Mill Power, Inlet mill Draught, Outlet Mill Temperature, Delta Pressure Mill, Delta Pressure Bag Filter, Mill Vibration, Residue 45µ and Blaine. Manipulated variables consisting of: Fresh Feed, Bag Filter Fan Speed and Separator Speed and variable output, namely quality. This study succeeded in developing a Machine Learning-based prediction model using the Random Forest Regressor (RFR) method which is integrated with the concept of Predictive Control Model (MPC) to predict the parameters of the cement milling process in VRM. The model was built using historical data of the production process that has gone through the preprocessing stage, including data cleaning, data integration, normalization, and validation.

The developed model is able to represent the relationship between process input variables and product quality output variables, especially residue parameters and Specific Energy Consumption (SEC). The Random Forest Regressor approach has been proven to be able to capture the nonlinear relationships between process variables quite well and has resistance to noise and variation of operation data. Based on the results of model validation using multiple optimizers, it was obtained that SLSQP optimizers provide better performance compared to without optimizers. The implementation of the Predictive Control Model (MPC) based on Random Forest Regressor (RFR) has the potential to increase the efficiency of cement milling operations by maintaining product quality according to blaine and residue targets, while reducing electrical energy consumption or Specific Energy Consumption (SEC). This approach is expected to be an alternative to conventional carrier-based control systems towards a more adaptive, stable, and efficient automatic control system. Quantitatively, the MPC-RFR model achieved R^2=0.99936, RMSE=11.062, MAE=2.488, MAPE=1.12%, validated by 5-fold cross-validation (R^2=0.9987±0.0003). SEC was reduced from 35.47 kWh/ton (conventional PID) to 29.46 kWh/ton with MPC-RFR (16.94% reduction), and further to 27.47 kWh/ton with MPC-RFR-SLSQP (22.55% reduction), corresponding to estimated annual savings of IDR 8.63 billion and IDR 11.48 billion respectively.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to Universitas Indonesia for its academic support throughout this research. The authors also appreciate the management and technical personnel of the cement manufacturing plant for providing operational data and valuable technical insights that enabled the development and validation of the proposed Machine Learning-based Model Predictive Control framework. Finally, the authors thank all colleagues and reviewers whose constructive comments contributed to improving the quality of this manuscript.

AUTHOR CONTRIBUTION STATEMENT

Elman Rudolf Pakpahan: Conceptualization, Methodology, Software, Data Curation, Formal Analysis, Investigation, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing. Iwa Garniwa: Supervision, Conceptualization, Methodology, Resources, Validation, Funding Acquisition (if applicable), Project Administration, Writing – Review & Editing.

REFERENCES

Ali, A., Anwar, S., & Izadian, A. (2025). Comparative Modeling of Vanadium Redox Flow Batteries Using Multiple Linear Regression and Random Forest Algorithms. *Energy Storage and Applications*, *2*(3), 11.

Atmaca, A., & Yumrutaş, R. (2014). Analysis of the parameters affecting energy consumption of a rotary kiln in cement industry. Applied Thermal Engineering, 66(1–2), 435–444. https://doi.org/10.1016/j.applthermaleng.2014.02.038

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Cui, K., Zheng, Y., Zhao, Y., Qin, Q., Liang, K., Chang, J., ... & Poon, C. S. (2025). Development of in-situ highly active calcium carbonate through anhydrous carbonation of OPC: Effect on hydration and properties of cement composites. *Cement and Concrete Research*, *197*, 107980. https://doi.org/10.1016/j.cemconres.2025.107980

Ghalandari, V., Esmaeilpour, M., Payvar, N., & Reza, M. T. (2021). A case study on energy and exergy analyses for an industrial-scale vertical roller mill assisted grinding in cement plant. *Advanced Powder Technology*, *32*(2), 480-491. https://doi.org/10.1016/j.apt.2020.12.027

Kumar, S., Gangotra, A., & Barnard, M. (2025). Towards a net zero cement: strategic policies and systems thinking for a low-carbon future. *Current Sustainable/Renewable Energy Reports*, *12*(1), 5. https://doi.org/10.1007/s40518-025-00255-2

Lepioufle, J. M., Marsteen, L., & Johnsrud, M. (2021). Error prediction of air quality at monitoring stations using random forest in a total error framework. *Sensors*, *21*(6), 2160. https://doi.org/10.3390/s21062160

Liu, R., Yu, J., Liu, L., Wang, Z., Zhou, S., & Zhu, Z. (2025). A Cement Bond Quality Prediction Method Based on a Wide and Deep Neural Network Incorporating Embedded Domain Knowledge. *Applied Sciences*, *15*(10), 5493. https://doi.org/10.3390/app15105493

McElroy, P. D., Bibang, H., Emadi, H., Kocoglu, Y., Hussain, A., & Watson, M. C. (2021). Artificial neural network (ANN) approach to predict unconfined compressive strength (UCS) of oil and gas well cement reinforced with nanoparticles. *Journal of Natural Gas Science and Engineering*, *88*, 103816.

Panagoda, S. S. (2023). Cement manufacturing process and its environmental impact. *J. Res. Technol. Eng*.

Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, *9*(3), e1301. https://doi.org/10.1002/widm.1301

Pönisch, D. L., Bittig, H. C., Kolbe, M., Schuffenhauer, I., Otto, S., Holtermann, P., ... & Rehder, G. (2025). Variability of CO2 and CH4 in a coastal peatland rewetted with brackish water from the Baltic Sea derived from autonomous high-resolution measurements. *Biogeosciences*, *22*(14), 3583-3614. https://doi.org/10.5194/bg-22-3583-2025

Saedi, M., Behfarnia, K., & Soltanian, H. (2019). The effect of the blaine fineness on the mechanical properties of the alkali-activated slag cement. Journal of Building Engineering, 26, 100897. https://doi.org/10.1016/j.jobe.2019.100897

Sahoo, N., & Kumar, A. (2022). Review on energy conservation and emission reduction approaches for cement industry. *Environmental Development*, *44*, 100767. https://doi.org/10.1016/j.envdev.2022.100767

Soomro, M., Tam, V. W., & Evangelista, A. C. J. (2023). Production of cement and its environmental impact. In *Recycled Concrete* (pp. 11-46). Woodhead Publishing.

https://doi.org/10.1016/B978-0-323-90407-9.00008-4

Sverdrup, H. U., & Olafsdottir, A. H. (2023). Dynamical modelling of the global cement production and supply system, assessing climate impacts of different future scenarios. *Water, Air, & Soil Pollution*, *234*(3), 191. https://doi.org/10.1007/s11270-023-06187-5

Tong, R., Sui, T., Feng, L., & Lin, L. (2023). The digitization work of cement plant in China. *Cement and Concrete Research*, *173*, 107266. https://doi.org/10.1016/j.cemconres.2023.107266

Yang, J., Zhao, H., He, X., Zeng, J., Su, Y., Li, Y., ... & Tian, C. (2024). Achieving carbon utilization and storage (CUS) in cement-based materials with wet-grinding carbonated concrete slurry waste. *Cement and Concrete Composites*, *152*, 105642. https://doi.org/10.1016/j.cemconcomp.2024.105642

Yang, S., Wan, M. P., Chen, W., Ng, B. F., & Dubey, S. (2020). Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization. Applied Energy, 271, 115147. https://doi.org/10.1016/j.apenergy.2020.115147

Zajac, M., Bolte, G., Skocek, J., & Haha, M. B. (2021). Modelling the effect of the cement components fineness on performance and environmental impact of composite cements. Construction and Building Materials, 276, 122108. https://doi.org/10.1016/j.conbuildmat.2021.122108

Zhang, K., Wang, J., Xin, X., Li, X., Sun, C., Huang, J., & Kong, W. (2022). A survey on learning-based model predictive control: Toward path tracking control of mobile platforms. *Applied Sciences*, *12*(4), 1995. https://doi.org/10.3390/app12041995