Machine Learning and Time Series Data Used in the Prediction and Regulation Model of Alumina Hydroxide Particle Size Distribution

Long Duan[1], Lijuan Qi[2], Yanfang Zhang[3], Shuai Shao[4], Qiaoyun Liu[5] and Linyu Li[6]
1, 4, 5, 6. Engineers
2, 3. Professor Level Senior Engineers
Zhengzhou Non-ferrous Metals Research Institute of Chalco (ZRI), Zhengzhou, China
Corresponding author: yanfang_zhang@chinalco.com.cn
https://doi.org/10.71659/icsoba2025-aa035

Abstract

Particle size control in the precipitation process of alumina production directly affects product quality and energy efficiency. To address challenges such as strong lag and large particle size fluctuations in traditional manual experience-based control, this study proposes a particle size distribution (PSD) prediction model and a dynamic control model based on machine learning and time series data. By integrating multi-source time series data, including equipment parameters, material characteristics, and process conditions, a high-dimensional dynamic dataset is constructed. Leveraging the multi-scale features of production data, a particle size distribution prediction model based on a CNN-LSTM (Convolutional Neural Network – Long Short-Term Memory) time series network is developed. A feature attention mechanism is introduced to enhance sensitivity to key process parameters, effectively capturing the non-linear coupling relationship between particle size variations and process parameters, thereby achieving prediction of PSD. Based on the prediction results, an optimized control model is established, which derives the control ranges of key parameters through reverse deduction, ultimately realizing stable control of particle size during the precipitation process. In this study, the average accuracy of PSD prediction exceeds 95 %, and particle size fluctuations are significantly reduced after control. This study not only verifies the effectiveness of time series data feature mining and machine learning modelling in process industries but also offers a "prediction-diagnosis-control" closed-loop paradigm for the intelligent optimization of high-dimensional coupled production systems, providing valuable guidance for the intelligent upgrading of alumina production.

Keywords: Seed precipitation, Particle size analysis, Machine learning, Prediction model, Production control.

1. Introduction

The Bayer process for alumina production offers the advantages of low energy consumption and high product quality. Seed precipitation is one of its key processes. By adding an appropriate amount of aluminium hydroxide $Al(OH)_3$, also known as alumina trihydrate ($Al_2O_3 \cdot 3H_2O$), or ATH) seeds to the supersaturated sodium aluminate liquor, alumina precipitates from the liquor in the form of ATH crystals through gradual cooling and continuous agitation. The particle size and strength of the alumina product largely depend on the particle size and morphological structure of ATH particles produced during the precipitation process. If the particle size of the ATH seeds in the seeded precipitation process is too fine, it can lead to poorer product quality; conversely, if the seeds are too coarse, the precipitation rate will decrease, which lowers the overall production yield [1, 2]. Therefore, accurately monitoring the seed PSD during the precipitation process and promptly adjusting production parameters are essential measures for stabilizing the seed PSD and improving product quality.

Due to the long process flow, numerous pieces of equipment, and complex material transformation in the precipitation process, key indicators such as the particle size of the produced

ATH are affected by multiple factors, including liquor composition, precipitation temperature setting, quantity and quality of seeds, precipitation time, impurity levels, and agitation intensity [1, 3]. Moreover, the process itself is characterized by large inertia, non-linearity, multiple disturbances, and complex correlations. As a result, relying solely on manual experiences and mechanism-based models fails to deliver timely and effective control strategies, leading to significant PSD fluctuations during the precipitation process.

With the advancement of automation and digitalization in alumina refineries, mature sensing devices can acquire equipment data in real time, while advanced detection technologies can periodically obtain material data from laboratory analyses. Effectively leveraging these vast amounts of production data makes it possible to extract key data features from complex data relationships, enabling the prediction of critical process indicators, and thereby guiding production control [4–8].

In this study, equipment and material data from the precipitation process in an alumina refinery were collected to extract time series information and construct a time series dataset. Based on machine learning methods, a PSD prediction model was developed, and optimization techniques were employed to identify the optimal feature. This approach ultimately guided production control, stabilized the PSD and reduced costs.

2. Data and Methods

2.1 Data Acquisition

The data used in this study were obtained from the Excellence Technology Center of Chalco and included equipment and material data from the precipitation process of three production lines over a period of three months (January to March 2025). The precipitation process consists of 14 precipitators, 7 heat exchangers, cyclone classifiers, and vertical disc filters, with each precipitation cycle of 44 hours. The equipment data cover measurements from the 14 precipitators and 7 heat exchangers, while the material data include PSD information of precipitation green liquor, precipitation spent liquor, and filter cake from the vertical disc filters. Detailed data are shown in Table 1.

In the seeded precipitation process, the precipitation green liquor flows sequentially from precipitator F1 through precipitators F2 to F14, with each precipitator equipped with agitation devices. The plate heat exchangers, labelled H1 to H7, are installed between precipitators F2 and F9 to progressively cool the sodium aluminate solution, thereby facilitating the precipitation of qualified ATH. A cyclone classifier is positioned above precipitator F11. Its underflow is directed to the horizontal disc filter, while the overflow and remaining slurry are returned to tank F12, re-entering the precipitation system. The slurry from precipitators F13 and F14 is collectively conveyed to the vertical disc filter. The resulting filter cake is fed into the seed tank to continue participating in the precipitation cycle, the filtrate proceeds to the evaporation stage, and the overflow is returned to precipitators F13 and F14.

Table 1. Detailed data.

Type	Sources	Item	Acquisition frequency
Equipment data	Precipitator	Temperature (°C)	Every hour
		Agitation current (A)	
		Liquid level (m)	
	Heat exchanger	Water temperature at outlet (°C)	
		Material temperature at outlet (°C)	
Material data	Precipitation green liquor / Precipitation spent liquor	α_k (Mass ratio of Na_2O_K to Al_2O_3)	Every 4 hours
		Suspended solids (g/L)	
		Al_2O_3 (g/L)	
		Fe_2O_3 (g/L)	
		SiO_2 (g/L)	
		Na_2O_K (g/L)	
		Na_2O_T (g/L)	
		Na_2O_C/Na_2O_T	
	Vertical disc filter cake	-14.24 μm particle size (%)	Once a day
		-29.4 μm particle size (%)	
		-44.49 μm particle size (%)	
		-59.16 μm particle size (%)	
		-80.72 μm particle size (%)	
		-101.9 μm particle size (%)	
		-125.3 μm particle size (%)	
		-150.2 μm particle size (%)	

2.2 CNN-LSTM Prediction Model

Convolutional Neural Networks (CNNs) have shown significant advantages over traditional shallow neural networks in feature extraction and selection, pattern recognition, predictive modelling, and classification tasks, thanks to their unique spatial feature learning mechanism. The CNN architecture comprises six key modules.

1) Input layer: standardizes and accepts multi-dimensional data inputs.
2) Convolutional layer: extracts local feature maps using learnable filters.
3) Activation function layer: introduces non-linear decision boundaries.
4) Pooling layer: performs down sampling to enhance feature invariance.
5) Fully connected layer: integrates high-level features.
6) Output layer: implements the final task mapping.

Compared with traditional Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks reduce the number of parameters by 40–60 % through a gated weight sharing strategy, while maintaining the capability to model long sequences with extended time steps. The architecture of an LSTM network comprises four core components.

1) Memory cell: a linear information channel that spans the entire sequence, accumulating information and preserving the continuity of long-term memory.
2) Forget gate: assesses historical information and control updates of the memory cell.
3) Input gate: filters new information and generates candidate values.
4) Output gate: regulates the data output.

The network architecture of the prediction model developed in this study is illustrated in Figure 1. It comprises a CNN feature extraction module and an LSTM prediction module. A parameter-free attention mechanism (SimAM) [9] is incorporated into the CNN module to enhance the effectiveness of feature extraction. The CNN module is used to extract features from the raw data, which are then fed into the LSTM module to capture time series information, ultimately achieving the output of PSD prediction.

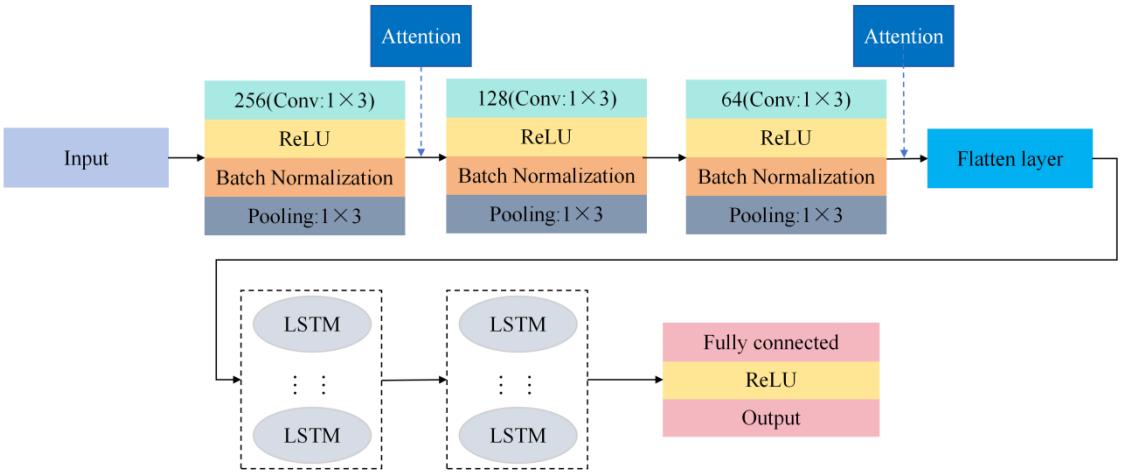

Figure 1. Architecture of the CNN-LSTM prediction model.

2.3 Control Model

The Genetic Algorithm (GA) is a metaheuristic optimization method inspired by the principles of natural selection and genetic inheritance. It encodes potential solutions as chromosomes mapped into a genetic search space and iteratively evolves the population according to the "survival of the fittest" principle. Through genetic operations such as selection, crossover, and mutation, GA efficiently searches the solution space. The core advantages of GA are as follows:

1) Global convergence: by combining probabilistic selection strategies with crossover and mutation operations, GA effectively avoids getting trapped in local optima.
2) Strong adaptability: it requires only a fitness evaluation derived from the objective function.
3) Parallel exploration mechanism: leveraging population-based synchronous evolution with chromosome encoding, GA achieves both broad coverage and in-depth search of the solution space.
4) Dynamic self-regulation: adaptive adjustment of crossover and mutation rates, combined with an elitism strategy, enables stable convergence of the algorithm.

The design of the GA-based control model developed in this study is illustrated in the flow chart shown in Figure 2. After inputting the data into the optimization model, key equipment data are adjusted and optimized to achieve the desired control objectives. Finally, an optimization strategy is generated to serve as a reference for production control. To evaluate the effectiveness of the optimization, the encoded key equipment data are merged with the complete time series data and then processed again by the prediction model to obtain the regulated and optimized predicted values of the PSD.

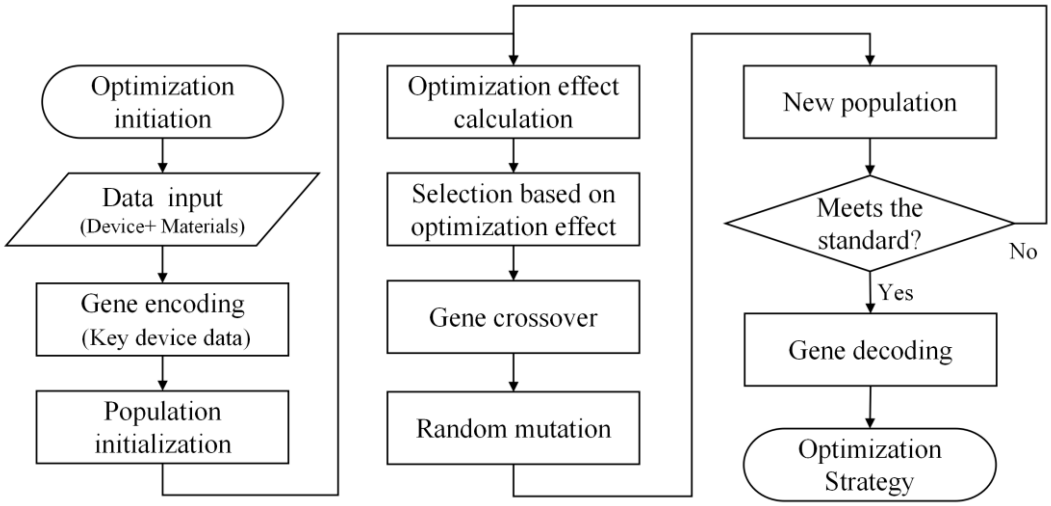

Figure 2. Flow chart of the optimization algorithm.

3. Experiment

3.1 Experimental Environment

The configuration environment used for all data processing and modelling in this study is summarized in Table 2.

Table 2. Configuration environment.

Configuration type	Configuration information
Operating system	Ubuntu 22.04
Programming language	Python
Development framework	PyTorch
Processor	Intel i9-13900K
Graphics card	NVIDIA GeForce RTX 4090 D (24GB)
Memory	64 GB

3.2 Data Preprocessing

The data collected during the experiment included a large amount of equipment data, some of these containing missing values. Before data modelling, all data were cleaned: only data items with complete time periods were retained. After data cleaning, each single time-node equipment dataset (collected once per hour) contained 38 data items; each single time-node material dataset (collected once every 4 hours) contained 13 data items; and the PSD dataset (collected once per day) contained 8 data items.

The fluctuations of key parameters during the precipitation process exhibited periodic patterns. This study aimed to construct datasets from massive time series production data, organized by the time sequence of each complete precipitation cycle (44 hours), in order to extract the relevant temporal and mechanistic information. As shown in Figure 3, each dataset contained 44 hours of equipment data and 10 material data nodes sampled every 4 hours (excluding the last data node), with a total of 1 672 equipment data items and 138 material data items. The datasets for the three production lines comprised 90, 84, and 87 data samples, respectively, which were divided into

training, validation, and test sets in a ratio of 3:1:1. The prediction targets were the PSD data at the next time node for particles smaller than 44.49 μm, 80.72 μm, and 101.9 μm.

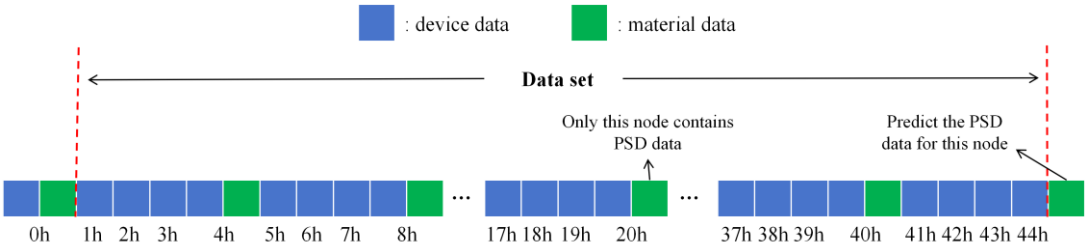

Figure 3. Dataset.

3.3 Establishment of Prediction Model

In establishing the prediction model, the Adaptive Moment Estimation (Adam) optimizer was used to update the model parameters, and the root mean square error was adopted as the loss function to optimize the model. The learning rate for training model was set as 1×10^{-4}, with the number of training rounds set as 200. Separate models were built for each of the three production lines. The accuracy (AC) was used to evaluate the prediction model, and AC was calculated according to Equation (1).

$$AC(G,G') = 1 - \sum_{i=1}^{n} \left|\frac{G_i - G'_i}{G_i}\right| \quad (1)$$

where:

G True value of particle size distribution
G' Predicted value of particle size distribution
n Quantity of data

3.4 Establishment of Control Model

In establishing the control model, the Genetic Algorithm (GA) was adopted with a population size set as 100, and the number of genes corresponded to the adjustable key parameters. The optimization ran for 100 iterations, and constraints were applied to the adjustable ranges of these parameters during the optimization process. The tuning target, that is the recent mean value of a specific particle size distribution, was used as the standard value of tuning the parameters. For adaptability calculation, the absolute error between the optimized value and the standard value was computed. After the optimization was completed, the volatility (standard deviation, SD) was used to evaluate the performance of the predictive control model, and the volatility was calculated according to Equation (2).

$$\mathrm{SD(G)} = \sqrt{\frac{\sum_{i=1}^{n}(G_i - \bar{G})^2}{n}} \quad (2)$$

where:

G True value of particle size distribution
n Quantity of data

4. Results and Discussion

4.1 Data Preprocessing Results

The particle size distributions of particles smaller than 44.49 µm, 80.72 µm, and 101.9 µm in the vertical disc filter cake during the precipitation process of the three production lines over a continuous three-month period are shown in Figure 4. The particle size distributions exhibited continuous fluctuations. The average proportion of particles smaller than 44.49 µm was 13.68 %, that of particles smaller than 80.72 µm was 62.97 %, and that of particles smaller than 101.9 µm was 86.07 %.

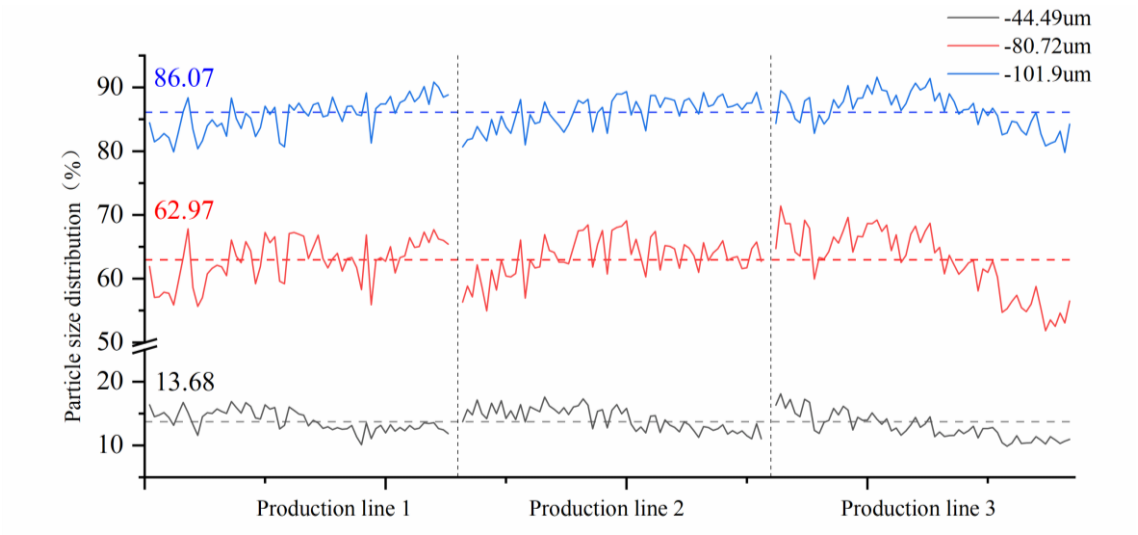

Figure 4. Particle size distribution.

4.2 Results from the Prediction Model

The prediction targets were the percentage shares of particles smaller than 44.49 µm, 80.72 µm, and 101.9 µm. During the modelling process, separate models were built for each of the three production lines. The precision for the test set and the modelling set is shown in Table 3, and the detection results for the test set are illustrated in Figures 5–7. For the three production lines, the prediction accuracy for each particle size exceeded 90 %, demonstrating the effectiveness of using machine learning methods to predict PSD and allowing for early forecasting of its variation.

Table 3. Relative errors of prediction models.

Production lines	Prediction accuracy (%)					
	-44.49 µm		-80.72 µm		-101.9 µm	
	Modelling	Test	Modelling	Test	Modelling	Test
Line 1	95.19	93.77	98.54	96.01	98.40	97.54
Line 2	94.98	93.91	97.23	96.77	99.03	97.97
Line 3	98.08	91.74	98.74	94.98	99.34	97.05

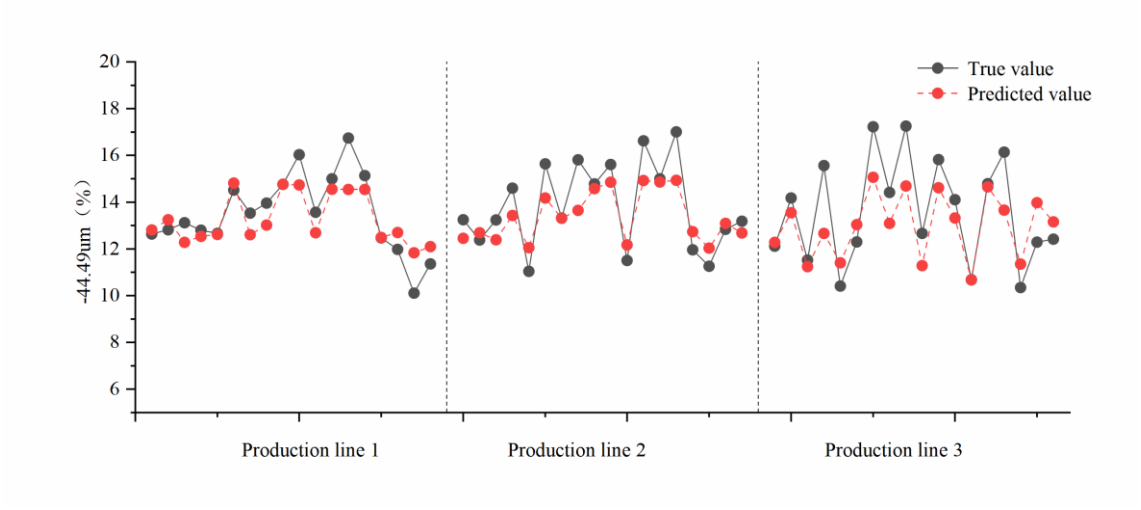

Figure 5. Prediction results of the proportion of particles smaller than 44.49 μm.

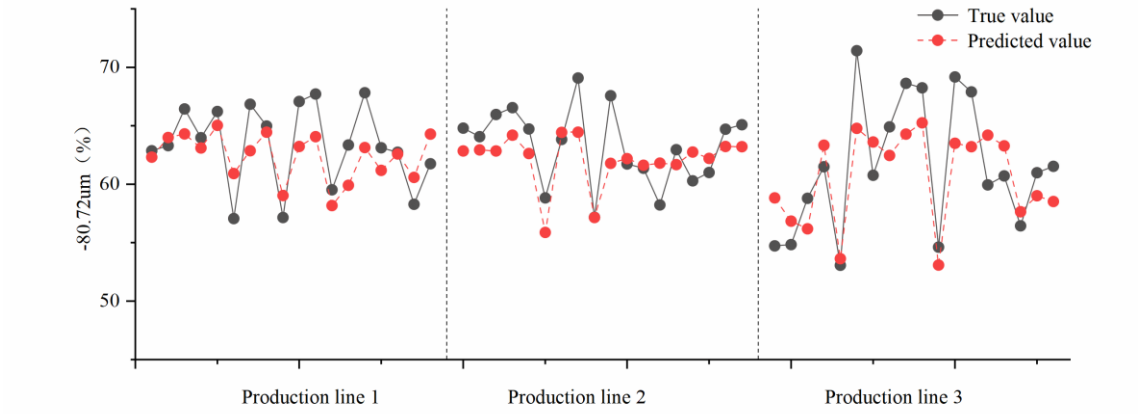

Figure 6. Prediction results of the proportion of particles smaller than 80.72 μm.

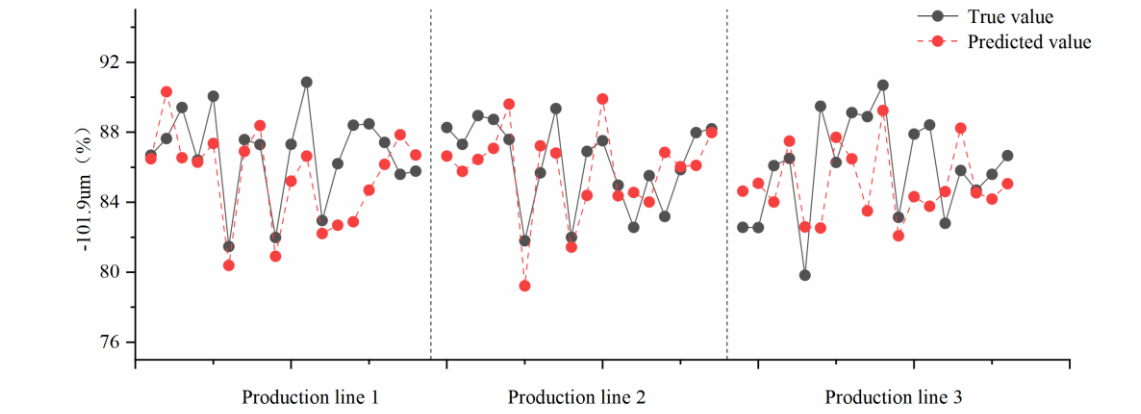

Figure 7. Prediction results of the proportion of particles smaller than 101.9 μm.

The test results of all production lines indicate that the prediction accuracy for smaller particle size distributions is lower than that for larger particle sizes. Therefore, in this study, feature visualization analysis was conducted on the prediction model for particles smaller than 44.49 μm. The results of feature visualization, shown in Figure 8 (using data from Production Line 2), reveal that the key sensitive features identified by the model included the historical PSD data from the previous day and the agitation intensity of each precipitator throughout the entire cycle. Meanwhile, the analysis indicates that the agitation intensity of the upstream precipitators exhibited higher sensitivity during the early stage of the precipitation cycle, whereas the agitation intensity of the downstream precipitators became more sensitive during the later stage of the precipitation cycle. These visualization results align well with the actual mechanistic phenomena observed in production, demonstrating the effectiveness of the prediction model.

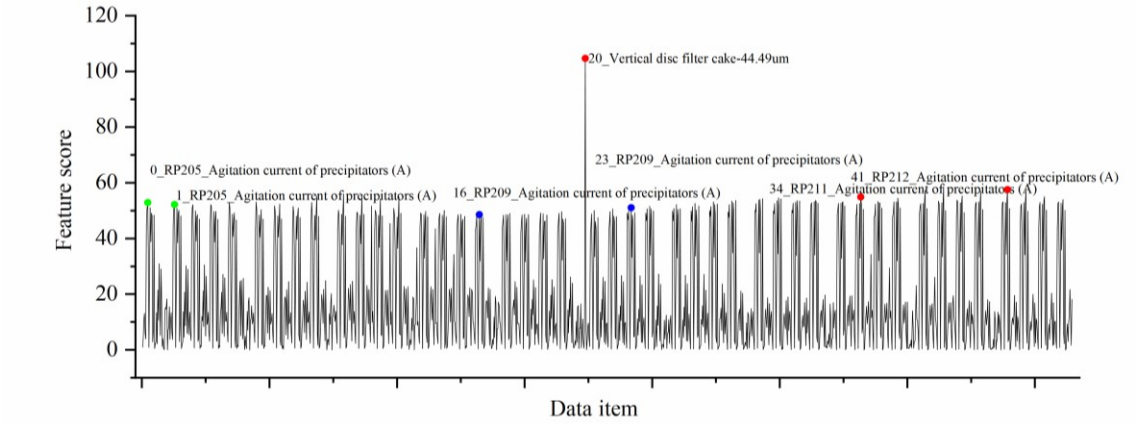

Figure 8. Feature visualization of the prediction model for particles smaller than 44.49 μm.

4.3 Result of Control Model

In the precipitation process, the percentage of particles smaller than 44.49 μm is an important indicator, and maintaining the stability of this PSD is of significant importance. Therefore, this indicator was first selected for optimization in constructing the control model, as described in Sections 2.3 and 3.4. For the control, 10 data samples that deviate from the average PSD value were selected as targets for optimization. In the adaptability calculation of the GA model, the standard parameter value was set as the mean proportion of particles smaller than 44.49 μm, i.e., 13.68 %. Based on the feature scores shown in Figure 8 and considering the controllable parameters of the precipitation process, the control parameters of the optimization algorithm were set as the more sensitive agitation currents of the precipitators. By adjusting the agitation currents, the agitation intensity can be regulated, thereby achieving optimization of the PSD during the precipitation process. In the process of gene encoding, the agitation currents for all time periods of all precipitators throughout the precipitation cycle were included in the encoding, resulting in an encoding length of 14 precipitators × 44 hours = 616. The results of the control and optimization are shown in Figure 9. For the 10 data samples, the standard deviation before optimization was 3.03, while after control, it was reduced to 0.60, with the PSD fluctuation significantly reduced. During the control and optimization process, the model provided optimization reference values for control parameters, namely the set agitation currents for each precipitator at each time period. Process personnel can refer to these values for production control.

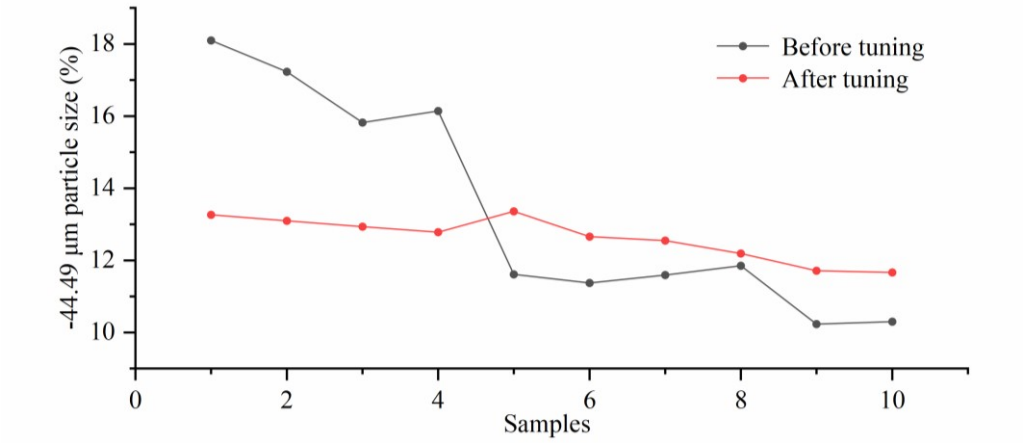

Figure 9. Tuning effect of the -44.49 μm particle size.

5. Conclusions

This study adopts machine learning techniques to establish a PSD prediction model and a control and optimization model based on production time series data, thereby realizing the early prediction of process indicators and the generation of production control strategies. The main contributions are as follows: Data on equipment and materials during the precipitation processes from three production lines over three consecutive months were collected; By integrating multi-source time series data, including equipment parameters, material characteristics, and process conditions, a high-dimensional dynamic dataset was constructed; By incorporating multi-scale features of production data, a PSD prediction model based on a CNN-LSTM time series network was developed. A feature attention mechanism was introduced to enhance sensitivity to key process parameters, effectively mapping the non-linear coupling relationships between particle size variation and process parameters. As a result, the model achieved prediction of the particle size distributions of -44.49 μm, -80.72 μm, and -101.9 μm, with accuracies of 91.74 %, 94.98 %, and 97.05 % respectively. Based on the prediction results, an optimization and control model was established. By reversely deriving the control ranges of key parameters, stable control of PSD in the precipitation process was achieved, reducing the standard deviation of -44.49 μm particle size fluctuations from 3.03 to 0.60. This study not only demonstrates the effectiveness of time series data feature extraction and machine learning modelling in process industries but also provides a "prediction-diagnosis-control" closed-loop paradigm for the intelligent optimization of high-dimensional coupled production systems, offering significant potential for promoting the intelligent upgrading of alumina production.

6. References

1. Wangxing Li, Theory and Technology of Alumina Production, Zhongnan University Press, 2010 (in Chinese).
2. Yusheng Wu, et al., Mechanism and patterns of particle size variation of aluminium hydroxide products during the precipitation process, *The Chinese Journal of Nonferrous Metals*, 2005, (12): 2060–2065 (in Chinese).
3. Chenglin Liu, et al., Optimizing seed-induced nucleation for enhanced Al(OH)3 crystal precipitation from supersaturated potassium aluminate solution, *Crystal Research and Technology*, 2024, 59(8): 2400086–2400086.
4. Long Chen, et al., Data-based review of prediction methods for production process indicators in the process industry, *Acta Automatica Sinica*, 2017, 43(06): 944–954 (in Chinese).

5. Kai G, et al., Data-driven prediction of quartz dissolution rates at near-neutral and alkaline environments, *Frontiers in Materials* 2022, 9, 924834.
6. Yongxiang L, et al., A self-supervised temporal temperature prediction method based on dilated contrastive learning, *Journal of Process Control* 2022, 120, 150–158.
7. Yufei Zhang, Laishi Li, Qi Yang, Prediction method for the caustic ratio in alumina digestion based on BP Neural Network, *Light Metals*, 2017, (09): 53–58 (in Chinese).
8. Duan Long, Shao Shuai, Zhang Yanfang, Predicting precipitation rate in alumina production using machine learning, *Proceedings of 42st International Conference of ICSOBA*, 27–31 October 2024, Lyon, France, *TRAVAUX* 53, 447–456.
9. Yang L, et al., SimAM: A Simple, parameter-free attention module for Convolutional Neural Network, PMLR, 2021.