Predicting Precipitation Rate in Alumina Production using Machine Learning

Long Duan[1], Shuai Shao[2] and Yanfang Zhang[3]
1. Assistant Engineer
2. Engineer
3. Professor
Zhengzhou Non-ferrous Metals Research Institute of CHALCO, Zhengzhou, China
Corresponding author: zyy_zyf@rilm.com.cn
https://doi.org/10.71659/icsoba2024-aa018

Abstract

The Bayer process for alumina production is complex and lengthy. Accurately and timely predicting process indicators and determining important parameters is the key to optimizing this process. This study focuses on the precipitation process during alumina production, achieving accurate and timely prediction of precipitation rate by using machine learning. Equipment and material data were collected over a 3-month period in the precipitation area of an alumina plant. The data was pre-processed and a time-series dataset was established with a total of 535 data samples (divided into training set, validation set, and testing set in a 3:1:1 ratio). Principal component analysis (PCA), partial least squares (PLS) and Pearson correlation analysis (Pearson) were used to select effective parameters. Four prediction models, including linear regression (LR), support vector machine (SVM), back propagation neural network (BP), and convolution neural network (CNN), were established based on effective parameters data. The coefficients of determination for the prediction model established in this study were all greater than 0.8 on the testing set. The results showed that by using mass production data to establish a machine learning model, we can accurately and timely predict the precipitation rate in the Bayer process. This provides a theoretical basis for predicting process indicators for the entire alumina production process and lays the foundation for the intelligent production of alumina.

Keywords: Alumina, Precipitation Rate, Machine Learning, Prediction Model.

1. Introduction

The Bayer process has advantages in low energy consumption and high product quality. In the Bayer process, the precipitation process is very important. Its main function is to precipitate aluminium hydroxide from the sodium aluminate solution via a sedimentation process. The goal of the precipitation process is to achieve the precipitation rate as close to the upper limit as possible while ensuring the qualified particle size of aluminium hydroxide. Precipitation efficiency directly affects the efficiency in production [1].

It is difficult to obtain the precipitation rate accurately and timely due to the complex and lengthy process, large number of devices and equipment, and the complex material transformation process of precipitation. Moreover, the process has the complexity of large inertia, nonlinearity, and many influencing factors and dependencies. Relying solely on human experience and mechanistic models cannot provide effective production control plans, resulting in low efficiency in the precipitation process.

With the progress of sensor technology and the continuous advancement of factory digitization, a large amount of production data can be obtained in the production process. Many scholars and process engineers have begun to study how to make full use of this production data, extract important values, predict key process indicators, and then guide production regulation [2–5].

In the aluminium smelting industry, there has been much research on prediction of process indicators for aluminium production by electrolysis [6-8]. However, in the Bayer industry, prediction of process indicators in alumina production mainly focus on the research of mechanistic and empirical models, with only few input parameters selected for modelling. The accuracy of these models is generally low, and the prediction results cannot guide real-time production regulation [9, 10].

Zhang et al. [11] constructed a time-series dataset using data from the first and last precipitation tank, and seed crystal size. Based on machine learning technology, this accurately and timely predicted the crystal seeds particle size attenuation. However, the device data and material data generated by each device in the whole production process are related to the process indicators. Using all available production data can make the predicted results to the actual production, but the increase of data items introduces more redundant information. Therefore, it is necessary to choose appropriate data processing methods and adopt efficient machine learning techniques to establish the prediction model.

In this study, the device data and material data of the precipitation process of an alumina plant were used to construct a time-series dataset, and a precipitation rate prediction model was established based on machine learning to guide production control, and ultimately improve efficiency in production with reduced costs.

2. Data and Methods

2.1 Data Acquisition

The data used in this study came from the Excellent Technology Centre of Chinalco, which supplied device and material data from the precipitation process in a production line for three months (September 2023 to November 2023). The precipitation process included 14 precipitation tanks, six heat exchangers, a classifying hydrocyclone, and a horizontal table filter. The precipitation period of the production line was 44 hours in a batch process.

The device data included data for 14 precipitation tanks and 6 heat exchangers. The material data included data for the precipitation raw liquid and the precipitation mother liquid. Table 1 shows the detailed data.

Table 1. Detailed data.

Data source	Data item	Data frequency
Precipitation tanks	Temperature (°C)	Every hour
	Stirring current (A)	
	Liquid level (m)	
Heat exchangers	Flow rate (m^3/h)	
	Feed pressure on heat source (MPa)	
	Feed pressure on cold source (MPa)	
	Outlet temperature of heat source (°C)	
	Outlet temperature of cold source (°C)	
Precipitation raw liquid & Precipitation mother liquid	α_k (Molar ratio of Na_2O_K to Al_2O_3)	Every four hours
	Suspended matter (g/L)	
	Al_2O_3 (g/L)	
	Na_2O_K (g/L)	
	Na_2O_T (total, g/L)	

5. Conclusion

In this study, device and material data of the precipitation process in a production line for three months (September 2023 to November 2023) was collected, and a time-series dataset containing 535 valid data samples was obtained through data pre-processing. PCA, PLS, and Pearson were used to reduce the data dimensionality of the dataset to obtain the modelling dataset. The key features (outlet temperature of heat exchanger cold source, α_K, and Na_2O_K) were obtained through analysis of the dimensionality reduction results. Based on the modelling dataset, LR, SVM, BP, and CNN precipitation rate prediction models were established. The greatest R^2 was 0.9288, obtained by CNN with PLS. The results showed that using mass production data to establish machine learning model can accurately and timely predict the precipitation rate in the Bayer process. It provides a theoretical basis for predicting the process indicators of the entire alumina production process and lays the foundation for the intelligent production of alumina.

6. References

1. Shao Shuai, Study on Influencing Factors of Precipitation Ratio, *Proceedings of 41st International Conference of ICSOBA*, 5–9 November 2023, Dubai, UAE, *TRAVAUX* 52, 623–630.
2. Jinliang Ding et al., Data-based multiple-model prediction of the production rate for hematite ore beneficiation process, *Control Engineering Practice* 2015, 45, 219-229.
3. F. Nakhaei et al., Recovery and grade accurate prediction of pilot plant flotation column concentrate: Neural network and statistical techniques, *International Journal of Mineral Processing* 2012, 110-111, 140-154.
4. Chen Long et al., Data-driven Prediction on Performance Indicators in Process Industry: A Survey, *Acta Automatica Sinica* 2017, 43(6), 944-954.
5. Kai Gong et al., Data-Driven Prediction of Quartz Dissolution Rates at Near-Neutral and Alkaline Environments, *Frontiers in Materials* 2022, 9, 924834.
6. Ye Zhu et al., Temperature prediction of aluminum reduction cell based on integration of dual attention LSTM for non-stationary sub-sequence and ARMA for stationary sub-sequences, *Control Engineering Practice* 2023, 138, 105567.
7. Xiaoxue Wan et al., A novel multiple temporal-spatial convolution network for anode current signals classifcation, *International Journal of Machine Learning and Cybernetics* 2022, 13, 3299-3310.
8. Yongxiang Lei et al., A self-supervised temporal temperature prediction method based on dilated contrastive learning, *Journal of Process Control* 2022, 120, 150-158.
9. Marija Savic et al., ANFIS-Based Prediction of the Decomposition of Sodium Aluminate Solutions in the Bayer Process, *Chemical Engineering Communications* 2016, 203(8), 1053-1061.
10. Mohammadreza Baghoolizadeh et al., Using of artificial neural networks and different evolutionary algorithms to predict the viscosity and thermal conductivity of silica-alumina-MWCN/water nanofluid, *Heliyon* 2024, 10(4), e26279
11. Zhang Yufei et al., Precast method of digestion RSA based on BP neural network, *Light Metals* 2017, 9, 53-58.