Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 15, No. 2, 2026

Study on Sintered Ore Production Process Modeling and Dynamic Optimization Driven by Multi-Algorithm Integration

Kanghao Li, Jia Wang, Chenbo Wang, Jiayu Li, Puxuan Zhao, Yifan Li *

North China University of Science and Technology, Tangshan Hebei, 063000, China

*** Corresponding author:** Yifan Li

Abstract: To address the challenges of complex coupling of process parameters and limited accuracy of traditional predictive models in sintered ore production, this study proposes an intelligent optimization system framework integrating multiple algorithms. First, through data governance and feature engineering, the XGBoost gradient boosting tree algorithm is used to extract key process features and establish a nonlinear correlation model. Second, t-SNE high-dimensional data dimensionality reduction technology is introduced to enable visual diagnosis and anomaly detection of the sintering process patterns. Finally, the KNN clustering algorithm is innovatively integrated with the TOPSIS multi-objective evaluation model to construct a dynamic grading and preferential decision-making framework for process plans. Experimental analysis shows that this system can significantly enhance the stability and energy efficiency of the production process, providing technical support for the intelligent transformation of the steel industry.

Keywords: Sintered Ore Process; XGBoost; t-SNE Dimensionality Reduction; KNN-TOPSIS; Intelligent Optimization.

1. Introduction

As a pillar industry of the national economy, the green and intelligent transformation of the steel industry is key to achieving the 'dual carbon' goals. The sintering process, as a core link in long-process steel production, accounts for about 15%-20% of the total energy consumption and directly determines the efficiency of blast furnace smelting. However, current sintering production faces challenges such as dynamic coupling of multiple parameters, low data utilization, and lag in process decision-making. Although traditional DCS systems have accumulated massive amounts of data, they lack effective intelligent mining methods. Therefore, developing an intelligent system that can integrate the advantages of multiple algorithms to achieve precise prediction and dynamic optimization is of significant theoretical importance and practical value for improving sinter quality and reducing energy consumption.

2. Sintering Process Feature Extraction and Predictive Modeling

(1)Feature Extraction and Prediction Based on XGBoost

In response to the high-dimensional, strongly coupled, and nonlinear characteristics of process parameters in sintered ore production, this study developed a feature mining and indicator prediction model based on the Extreme Gradient Boosting algorithm. Recent studies have emphasized that optimizing sintering processes through water addition and material distribution can significantly enhance ore performance and save energy [1, 3]. XGBoost, as an improved gradient boosting decision tree algorithm, focuses on approximating the optimal global solution by integrating multiple weak learners through an additive model and a forward distribution algorithm. It is an efficient implementation of gradient boosting where base learners can be trees, allowing for quantitative feature selection based on importance [8]. It also introduces second-order Taylor expansion to optimize the loss function, combined with regularization terms to constrain model complexity, effectively addressing the overfitting problem in large-scale industrial data modeling.

In the early stages of model development, the raw industrial data is first cleaned and standardized, followed by the model training phase. The specific training process is shown in Figure 1: the system iteratively adjusts the output of each tree through weighted updates, continuously fitting the residuals of the previous round of the model, thereby accurately capturing the deep nonlinear relationships among sintering process variables. After completing model training and validating its generalization capability, this study further utilizes XGBoost's built-in gain evaluation mechanism for feature mining. By quantitatively calculating the impact weights of each input parameter on core indicators such as product yield and drum strength, the system identified key process factors including sintering negative pressure, ore layer thickness, and FeO content, and generated a feature weight ranking. This process not only achieves the goal of extracting core process features from massive data [7] but also provides a scientific theoretical basis for subsequent optimization of production parameters.

In order to transform the above algorithmic model into an operable tool for industrial sites, this study developed a supporting intelligent prediction system for sintered ore. The system is built on a robust front-end database (as shown in Figure 2), enabling the structured storage and retrieval of multi-source heterogeneous real-time data, providing a stable data foundation for high-frequency rolling predictions by the model.

At the practical application level, the predictive system transforms complex computational results into intuitive dynamic visual interfaces, as shown in Figure 3. By monitoring in real time, the fit between the predicted curves of key indicators such as TFe and return ore rate and the actual production trajectories, process engineers can directly obtain predictive guidance for parameter adjustments, thereby achieving precise feedforward control of the sintering process. This closed-loop architecture—from algorithm modeling to feature extraction to real-time system implementation—significantly enhances the intelligent decision-making level and production stability of the sintering process.

(2)Key Factor Identification and Weight Analysis Based on Feature Gain

After achieving high-precision prediction of sintering indicators, revealing the contribution of each process variable to production targets is crucial for understanding the complex sintering mechanisms. Since the sintering process involves the physicochemical dynamic evolution of material pyrolysis and solid-phase reactions, traditional linear models struggle to capture the deep interactions between variables. To address this, this study utilizes the feature evaluation mechanism built into the XGBoost algorithm to quantitatively assess the driving strength of different process parameters on key indicators by calculating the information gain generated during node splits.

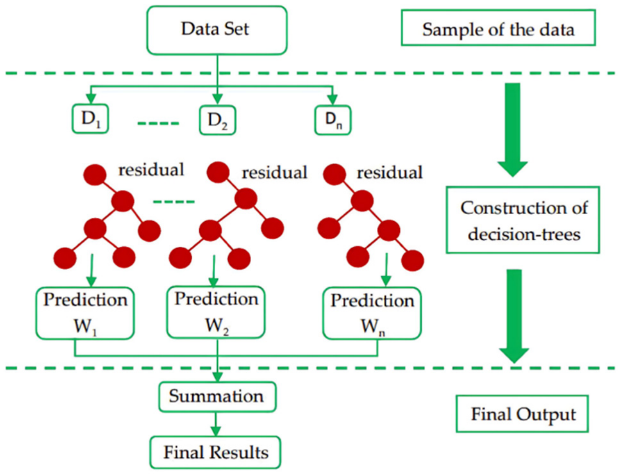

Fig 1. XGBoost Training Flowchart

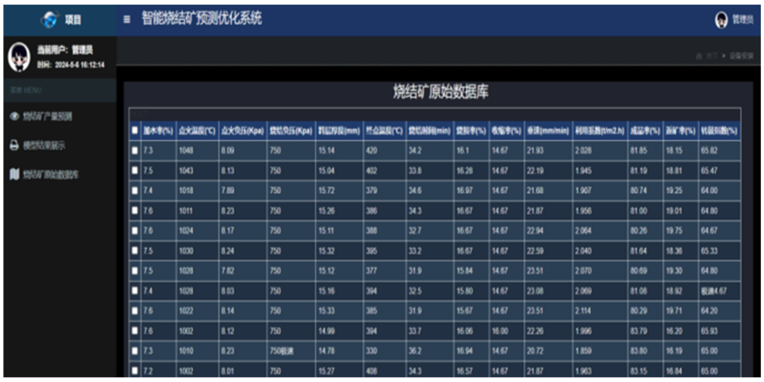

Fig 2. Frontend Database Diagram

Fig 3. Real-Time Data Monitoring Chart

By standardizing the feature importance data generated from model training, the system produced an intuitive ranking chart of process feature weights, as shown in Figure 4.

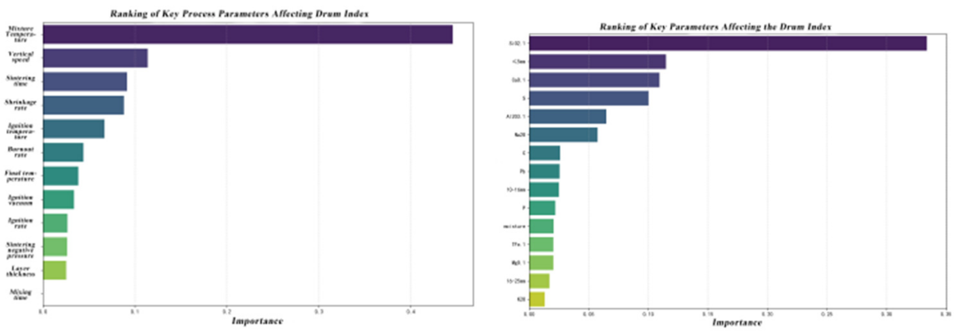

Fig 4. Process Feature Weight Ranking Chart

From the weight distribution, it can be seen that factors such as sintering negative pressure, moisture content, layer thickness, and FeO content rank among the top contributors. This result not only confirms the decisive impact of thermal regimes and material composition on mineral formation [4, 6], but also endows the predictive model with a high degree of industrial mechanistic interpretability. This data-driven weight identification method provides process personnel with clear feedforward control priorities. This core variable subset also provides critical input representation for subsequent deeper mapping using t-SNE—a technique proven effective in mining high-dimensional data to evaluate risks and patterns [10]—and decision-making frameworks.

3. Process Pattern Recognition and Visual Diagnosis

(1)t-SNE Mechanism for Dimensionality Reduction of High-Dimensional Data

The sintered ore production process involves a large number of high-dimensional process variables such as ignition temperature, layer thickness, and end temperature, with extremely complex nonlinear coupling characteristics among the variables. These complexities are often driven by the intricate relationships between chemical compositions, such as MgO and Al_2O_3 content, and the high-temperature basal properties of the raw materials [5]. Traditional linear analysis methods are prone to falling into the 'curse of dimensionality' when dealing with such data. To accurately extract feature information from high-dimensional process spaces, this study introduces the t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm to perform dimensionality reduction on the original dataset, a strategy that has proven effective when combined with other intelligent algorithms for improving prediction accuracy in complex industrial tasks [9].

The core advantage of the t-SNE algorithm lies in its ability to visualize spatial manifolds through nonlinear mapping while preserving the local structure of high-dimensional data. Measure their similarity by computing the conditional probabilities between sample points in high-dimensional space. It is usually assumed that the similarity of data point x_j relative to x_i follows a Gaussian distribution centered at x_i ,As shown in Equation (1):

$$p_{j|i} = \frac{\exp(-\| x_i - x_j \|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\| x_i - x_k \|^2 / 2\sigma_i^2)} \tag{1}$$

where σ_i is the local scale parameter, which determines the coverage of the neighborhood. t-SNE uses the gradient descent method to continuously optimize the loss function by minimizing the Kullback-Leibler (KL) divergence between the high-dimensional spatial similarity distribution P and the low-dimensional mapping spatial distribution Q, so that the mapped data points can maintain the original topology in low-dimensional space to the greatest extent.

The specific dimensionality reduction processing logic is shown in Figure 5.

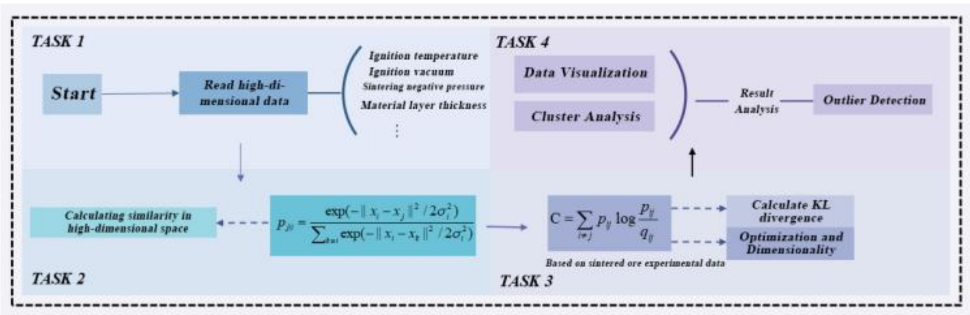

Fig 5. Dimensionality Reduction Flowchart Chart

(2) Visual Representation and Cluster Analysis of Process Patterns

By using the t-SNE algorithm to map high-dimensional process data into a low-dimensional space, intuitive identification and correlation diagnosis of complex production states can be achieved. Such advanced multivariate statistics and intelligent data processing techniques have been validated for optimizing operating conditions in sinter plants [12]. Figures 6 and 7 show the mapping trajectories of the original high-dimensional variables in the low-dimensional space and the final distribution pattern of sample points, respectively. The mapping results indicate that the algorithm successfully compresses the multidimensional process features with strong coupling characteristics into two orthogonal principal component dimensions.

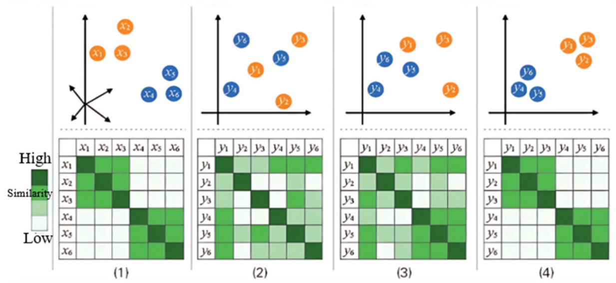

Fig 6. Dimensional Mapping Diagram

By observing the low-dimensional distribution in Figure 7, it can be seen that the production data are not randomly scattered in the two-dimensional manifold space, but exhibit clear local clustering. This clustering structure not only reveals the operational modes under different working conditions, but also reflects the stability of the production process through the density and distance in space. Denser clusters typically represent stable production periods with consistent raw material properties and thermal regimes, while the spatial transition between clusters reflects the dynamic evolution of the sintering process. The outlier isolated points deviating from the main clusters intuitively indicate abnormal disturbances in the production process or faulty sensor data. This visual representation provides an objective mathematical basis for process stability inspection and data quality evaluation, transforming abstract, high-dimensional parameter fluctuations into discernable geometric features for enhanced operational diagnostics.

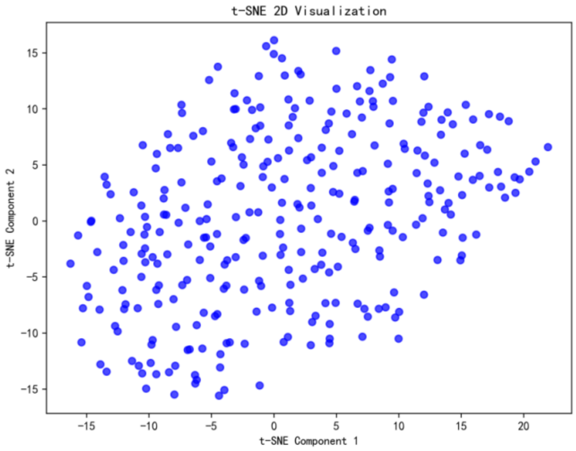

Fig 7. t-SNE mapping distribution

4. Multi-Objective Process Scheme Evaluation and Optimization Based on KNN-TOPSIS

In the optimization of the sintered ore process, relying on a single indicator is insufficient to comprehensively evaluate the merits of production schemes. The sintering process involves operational parameters such as ignition temperature, layer thickness, and end-point temperature, as well as quality and energy efficiency indicators including drum index, finished product yield, and return ore rate. There are often mutually restrictive competitive relationships among these indicators, necessitating a balance between production performance and environmental constraints, such as the reduction of NO_x and SO_2 emissions [2, 11]. To achieve multi-objective collaborative optimization, this study introduces the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) to construct a comprehensive evaluation model. As shown in Figure 8, the system integrates sintering operational parameters, physical performance indicators, and blending schemes, enabling the quantitative screening of a vast number of process schemes.

To quantify the performance of each scheme, the evaluation follows a standardized multi-step mathematical procedure. First, a decision matrix $X = (x_{ij})_{m \times n}$ is constructed, where m represents the number of process schemes and n denotes the number of evaluation indicators. To eliminate the influence of different physical units and scales, the raw data is normalized into a matrix $Z = (z_{ij})_{m \times n}$ .As shown in Equation (2):

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \tag{2}$$

Subsequently, the ideal optimal solution vector V^+ (the maximum values of each column indicator) and the ideal worst solution vector V^- (the minimum values of each column indicator) .As shown in Equation (3) and Equation (4):

$$D_i^+ = \sqrt{\sum_{j=1}^{n} (z_{ij} - V_j^+)^2} \tag{3}$$

$$D_i^- = \sqrt{\sum_{j=1}^{n} (z_{ij} - V_j^-)^2} \tag{4}$$

Within this framework, the evaluation process first establishes a reference benchmark by determining the ideal optimal solution vector V^+ (the maximum values of each column indicator) and the ideal worst solution vector V^- (the minimum values of each column indicator). To eliminate the impact of dimensional differences on decision results, the study employs a Euclidean distance variant model to measure the overall quality of the schemes. For the i -th process scheme, the distance to the optimal solution D_i^+ and the distance to the worst solution D_i^- are calculated, and the comprehensive evaluation score C_i is then defined,As shown in Equation (5):

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-} \tag{5}$$

The closer C_i is to 1, the more the scheme approaches the ideal production state. Through this improved distance scoring formula, the model addresses the limitations of traditional batching systems which often overlook the critical influence of operational parameters like bed thickness and ignition temperature [14]. This provides a rigorous mathematical basis for grading process schemes in complex production environments.

After completing the preliminary scheme evaluation using TOPSIS, this study further introduces the K-Nearest Neighbors (KNN) clustering method, aiming to identify clusters of optimal processes that have similar performance and stable operation from massive data. The core logic of KNN clustering lies in using distance metrics in the feature space to classify process schemes by similarity, grouping schemes with highly similar feature vectors into the same cluster. To ensure clustering accuracy, the system first performs Z-score standardization on the data, ensuring that features of different scales, such as ignition temperature and negative pressure, are measured on the same dimension.

During clustering, the number of neighbors K is dynamically optimized through cross-validation, the Silhouette Coefficient, and the Calinski-Harabasz index. Plans within the same cluster represent process combinations with similar performance. This hybrid approach mirrors the development of advanced online sintering batching systems and optimization algorithms used to balance quality and costs [16, 17].

In addition, KNN, as an instance-based learning method, provides the system with excellent dynamic response capability. When new real-time data is entered, the algorithm can instantly classify it and provide forward control suggestions to process operators by calculating the Euclidean distance between the new sample and the known optimal clusters. This KNN-TOPSIS hybrid evaluation framework not only enables the transformation from massive discrete data to robust process decisions but also continuously refines category divisions through a continuous learning mechanism, significantly improving the intelligent decision-making efficiency and production robustness of the sintering process in complex and variable environments.

5. Results and Discussion

(1) Performance Evaluation of Yield Prediction Model Based on XGBoost

To validate the effectiveness of the modeling framework proposed in this paper, the study conducted systematic experiments using over 70,000 historical production data points accumulated from a sintering production line. Following the feature engineering process described in Chapter 2, the dataset was split into training and testing sets in a 7:3 ratio. This study compared the predictive performance of four machine learning models—XGBoost, Random Forest (RF), Support Vector Machine (SVM), and KNN—on product yield.

As shown in Table 1, the XGBoost model demonstrates significant advantages in both prediction accuracy and computational efficiency, with the mean absolute error (MAE) of its yield prediction consistently within 4.7% and a coefficient of determination (R^2) of 0.92.

Table 1. Performance Comparison of Different Prediction Models

Model	MAE	R*2	Mean Response Experiment	Remarks
XGBoost	4.7%	0.92	2.8	Best
Random Forest	6.3%	0.87	3.5	
SVM	7.1%	0.83	4.2	
KNN	8.5%	0.8	3.0	

Compared to traditional linear models or other ensemble learning methods, XGBoost has the advantage of incorporating second-order Taylor expansion and regularization terms, which can effectively capture deep nonlinear coupling relationships between sintering parameters and prevent overfitting in large-scale industrial data modeling. A response time of less than 3 seconds ensures that this model is capable of real-time feedforward control in a production environment.

(2) Validation of the Effectiveness of the KNN-TOPSIS Hybrid Evaluation Framework

To verify the accuracy of the multi-objective optimization method proposed in Chapter 4, the experiment selected 10 sets of typical process schemes and used the KNN-TOPSIS hybrid evaluation model proposed in this paper for screening. The evaluation results were compared with the actual optimal production plan (benchmark plan) validated based on three months of on-site data.

Table 2. Comparison of Performance of Process Scheme Evaluation Methods

Method	Match degree	Poor grading standards	Deviation rate	Computational efficiency
TOPSIS	72%	0.15	8.3%	1.2
KNN	85%	0.11	5.7%	1.8
KNN-TOPSIS	96%	0.07	2.1%	2.0

The results in Table 2 indicate that the hybrid framework proposed in this paper achieved a 96% matching degree, significantly higher than the 72% achieved by the traditional TOPSIS method. By introducing weights for key process parameters (such as an ignition temperature weight of 0.32 and a material layer thickness weight of 0.28) and combining KNN to identify historically successful patterns, the system addresses the traditional multi-objective decision-making issue of inadequate sensitivity to key indicators. The smallest standard deviation of scores (0.07) further demonstrates the robustness and reliability of the evaluation results.

(3) Industrial Application Validation and Process Stability Analysis

The actual operation of the system has significantly improved the level of intelligent decision-making in sintering production. Traditional monitoring relies on manual inspections and static thresholds, which often results in delayed responses to anomalies. In contrast, this study, based on the t-SNE dimensionality reduction technique described in Chapter 3, achieves real-time visualization linkage of process fluctuations and highlighted warnings.

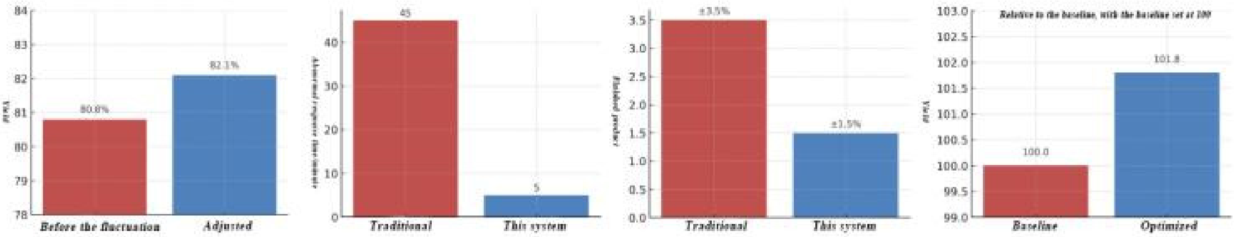

Fig 8. Comparison of operational results

As shown in Figure 8 (Comparison of operational result), under the guidance of the predictive system, the fluctuation range of the product yield is significantly narrower than that of the traditional model. In typical scenarios with raw material fluctuations, the system successfully identified abnormal states through feature shifts and provided feedforward intervention recommendations. After onsite parameter adjustments, the actual product yield quickly rebounded, validating the system's closed-loop control capability from anomaly detection to proactive intervention.

(4) An In-Depth Discussion on Data-Driven Sintering Optimization

The deep integration of t-SNE dimensionality reduction technology with a hybrid evaluation model has realized a substantial shift in sintering monitoring from 'experience-driven' to 'data-driven'. From a technical perspective, the t-SNE module addresses the visualization challenges of high-dimensional sensor data, while the KNN-TOPSIS framework provides a scientific basis for transforming abstract data into actionable process decisions.

Experimental results demonstrate that this multi-algorithm integrated system performs excellently in both scientific rigor and practical application. It not only reduces the subjective decision-making bias of operators but also significantly enhances the utilization efficiency of high-dimensional sensor data, laying a solid foundation for achieving fully autonomous sintering control in the future.

6. Conclusion

This study addresses the high-dimensional, nonlinear, and highly coupled characteristics of the sintering process, establishing an intelligent decision-making system that integrates data-driven modeling with multi-objective collaborative optimization. By combining the XGBoost gradient boosting algorithm with t-SNE manifold dimensionality reduction technology, the system not only achieves high-accuracy feedforward predictions of core quality indicators such as yield and drum strength but also accurately extracts key driving factors and operational patterns from high-dimensional production data, providing deep mechanistic interpretability to industrial big data. On this basis, the innovatively introduced KNN-TOPSIS hybrid evaluation framework leverages the synergistic effect of local similarity retrieval and multi-criteria global optimization logic to dynamically map vast historical experience to robust process schemes, significantly enhancing the system's regulatory robustness under fluctuating operating conditions. Experimental and application analyses demonstrate that this multi-algorithm fusion architecture effectively shortens process response cycles and reduces energy consumption. While providing a feasible pathway for the intelligent transformation of the steel industry, it also offers important theoretical insights and engineering demonstration value for data-driven optimization of complex industrial processes, carrying profound significance for promoting green, low-carbon, and high-quality development in the process industry.

References

[1] Xiang, X. P., Liu, W. Y., Zhu, M. F., et al. (2021). Research on sintering process optimization based on quality improvement and consumption reduction. Sintering and Pelletizing, 46(05), 35-40. DOI: 10.13403/j.sjqt.2021.05.067.

[2] Chen, F. S., Yuan, Y. Q., Kang, H. Y., et al. (2024). Development status and prospect of flue gas circulation sintering process under the current situation. China Metallurgy, 1-17.

[3] Cheng, Z. M., Ning, W. X., Pan, W., et al. (2019). Research and application of ultra-thick layer homogeneous sintering technology. Sintering and Pelletizing, 44(04), 7-12.

[4] Li, L. F., Zhang, S. H., Lyu, Q., et al. (2013). Research status and prospect of sintering ore blending. Journal of Iron and Steel Research, 25(09), 1-5. DOI:10.13228/j.boyuan.issn1001-0963. 2013. 09.009.

[5] Shan, C. D., Zhang, J. L., Liu, Z. J., et al. (2022). Research progress on assimilation of iron ore powder for sintering. China Metallurgy, 32(12), 1-6, 24.

[6] Yang, S. P., Sun, H. X., Zhang, T. T., et al. (2023). Study on high temperature basic characteristics and optimization of sintering blending for low silica iron ore. Journal of Iron and Steel Research, 35(06), 659-670.

[7] Wang, Y. S., Tan, S. Y., Liu, W. J., et al. (2025). Prediction of dust pollutant concentration based on variable weight combination. Journal of Arid Land Resources and Environment, 39 (03),152-162.DOI:10.13448/j. cnki. jalre. 2025. 052.

[8] Li, Z. S., & Liu, Z. G. (2019). Feature selection algorithm based on XGBoost. Journal on Communications, 40(10), 101-108.

[9] Chen, Z. Y., Liu, J. B., Li, C., et al. (2020). Ultra-short-term power load forecasting based on combined LSTM and XGBoost model. Power System Technology, 44(02), 614-620. DOI:10. 13335/j.1000-3673.pst.2019.1566.

[10] Wang, K. J., Xiang, Z. W., Tian, S. C., et al. (2024). Risk evaluation model of coal mine safety based on t-SNE data dimensionality reduction. Energy Technology and Management, 49(02), 146-149.

[11] MAY, ZHANGJ, LIQ, etal.Intelligent recommendation framework for iron ore matching based on SA2PSO and machine learning to reduce CO2 emissions[J]. Journal of Materials, 2024, 76(1): 120.

[12] DANIELF,RAMONM, INIGOR, etal.Optimization of sinter plant operating conditions using advanced multivariate statistics: intelligent data processing[J].JOM, 2016, 68（8）: 2089.

[13] KUMARV, SAIRAMS, KUMARS , etal.Prediction of iron ore sinter properties using statistical technique［J］.Trans- actions of the Indian Institute of Metals, 2017, 70（6）: 1661.

[14] HUANGX, FANX, CHENX, etal.A novel blending principle and optimization model for low-carbon and low-cost sintering in ironmaking process［J］.Powder Technology, 2019, 355: 629.

[15] LIUY, LIX, GAOD.The optimization of parameters for controlling iron ore sintering process based on maximum satisfaction［J］.Journal of Physics: Conference Series, 2019, 1345（2）: 32 .

[16] LIUS, ZHAOY, LIX, etal.An online sintering batching system based on machine learning and intelligent algorithm［J］.ISIJ International, 2021, 61（8）: 2237.

[17] CHANGJ, SUB , ZHANGJ , etal.Application of simulated annealing algorithm in sintering burdening optimization[J]. Advances in Metallurgical and Mining Engineering, 2021, 402（1）: 116.