AIMS Mathematics, 10(6): 13104–13129.
DOI: 10.3934/math.2025588
Received: 31 December 2024
Revised: 15 May 2025
Accepted: 20 May 2025
Published: 06 June 2025

https://www.aimspress.com/journal/Math

Research article

Application of a hybrid nonlinear algorithm driven by machine learning and feature importance identification for temperature control prediction of the bath smelting process

Senyuan Yang[1,2], Bo Yu[1,2], Jianxin Pan[3,4], Wuliang Yin[5], Hua Wang[1,2], Kai Yang[1,2,*] and Qingtai Xiao[1,2,*]

[1] State Key Laboratory of Complex Nonferrous Metal Resources Clean Utilization, Kunming University of Science and Technology, Kunming 650093, China
[2] Faculty of Metallurgical and Energy Engineering, Kunming University of Science and Technology, Kunming 650093, China
[3] Faculty of Science and Technology, Beijing Normal – Hong Kong Baptist University, Zhuhai, Guangdong 519087, China
[4] Research Center for Mathematics, Advanced Institute of Natural Sciences, Beijing Normal University, Zhuhai, Guangdong 519087, China
[5] School of Electrical and Electronic Engineering, The University of Manchester, Manchester M13 9PL, UK

* **Correspondence:** Email: kaiyang95@163.com; qingtai.xiao@kust.edu.cn.

Abstract: Temperature control in bath smelting processes is crucial for optimizing the efficiency and quality of metal extraction, especially for nickel and copper. Traditional prediction methods often fail to account for the nonlinear and complex nature of these processes. This work introduces a novel hybrid nonlinear analysis algorithm combining the random forest–least squares support vector machine (RF-LSSVM) and random forest–relevance vector machine (RF-RVM) models to enhance the accuracy of temperature prediction. Utilizing 868 datasets collected from an oxygen-enriched top-blown furnace, key parameters such as the feeding amount (X_1), oxygen pressure (X_2), oxygen flow (X_3), total air flow (X_7), and lance windpipe back pressure (X_5) were analyzed. The RF-LSSVM model achieved superior predictive performance, with a mean absolute error (MAE) of 7.58 and a root mean square error (RMSE) of 9.82 for matte temperature (Y_1), and an MAE of 10.47 and an RMSE of 13.31 for slag temperature (Y_2). Comparatively, traditional methods showed higher errors,

with MAE values of up to 23.64 and RMSE values as high as 59.14 in some cases. Additionally, the RF-RVM model performed significantly better than conventional models, with MAE and RMSE improvements of approximately 10~20%. These results demonstrate that the hybrid models effectively capture the intricate dynamics of the smelting process, offering a robust and adaptive framework for real-time temperature prediction. The improved accuracy in temperature control leads to enhanced smelting efficiency, reduced energy consumption, and higher quality of the extracted metals, ultimately benefiting the metallurgical industry by enabling more precise and sustainable production processes.

Keywords: bath smelting; temperature prediction; machine learning; hybrid models; oxygen-enriched top-blown furnace
Mathematics Subject Classification: 62M20, 90C26, 92B20

1. Introduction

Temperature control in the bath smelting process is a critical factor that influences the efficiency and quality of metal extraction, particularly for nickel and copper. Accurate temperature prediction and control are essential for optimizing the smelting process, enhancing product quality, and reducing operational costs. The inherent complexity and nonlinearity of the smelting process, influenced by multiple interacting factors such as feeding amount, oxygen pressure, oxygen flow, and pipeline pressure, necessitate advanced predictive models to manage these variables effectively [1–3]. Traditional empirical methods often fall short in capturing the intricate dynamics of the process, underscoring the need for innovative approaches. Furthermore, the environmental and economic pressures on the metallurgical industry require more efficient and sustainable practices, making precise control strategies even more crucial. Recent technological advancements in data acquisition and sensor technology have facilitated the collection of extensive process data, providing an opportunity to develop more sophisticated predictive models.

Traditional methods for temperature prediction in bath smelting processes, such as empirical correlations and basic statistical models, are subject to significant limitations due to their inability to account for complex process dynamics and nonlinear relationships. These methods are generally based on historical data and simplified assumptions that may not accurately reflect the complex and dynamic nature of the smelting environment [4–6]. Furthermore, the high costs and time-consuming nature of the experimental setups further restrict the applicability of these traditional methods in real-time industrial settings [7–9]. The reliance on linear assumptions and the inability to handle large datasets limit their accuracy and scalability, making them less suitable for modern, high-efficiency smelting operations [10–12]. Therefore, there is an urgent need to develop more robust and adaptable models capable of accurately predicting temperature fluctuations during the smelting process.

The advent of machine learning (ML) and advanced data analytics offers promising solutions to overcome the limitations of traditional methods. Machine learning algorithms, particularly those based on deep learning, have demonstrated remarkable capabilities in modeling complex, nonlinear systems by leveraging large datasets to uncover hidden patterns and relationships [13–16]. These algorithms are well-suited for applications in industrial process control, where they can provide real-time predictions and adaptive control strategies [17–19]. The ability of machine learning models to learn from historical data and improve over time makes them particularly useful in dynamic and

complex environments like smelting processes. Moreover, the integration of ML with internet of things (IoT) devices has further enhanced real-time monitoring and control capabilities. In particular, hybrid models that combine different machine learning techniques, such as the least squares support vector machine (LSSVM) and relevance vector machine (RVM), have shown superior performance in capturing the intricate dynamics of industrial processes [20–23].

The enhanced predictive accuracy and increased robustness of these hybrid models renders them more suitable for industrial applications. The continuous evolution of these models through ongoing research and development ensures their relevance and effectiveness in addressing emerging challenges in the smelting industry. Recent studies have highlighted the effectiveness of hybrid machine learning models in various industrial applications. For instance, deep learning models have been used to predict the formation energies and phase stability of perovskite oxides, showing promising accuracy compared with traditional methods [23,24]. Yang et al. (2024) proposed a hybrid model based on the radial basis function for predicting the surface fluctuation of slag cover, demonstrating improved predictive performance in smelting environments [25]. Assareh et al. (2023) employed a multi-objective evolutionary algorithm combined with machine learning to predict the phase changes in smelting processes, which proved effective in various industrial scenarios [26]. Hareharen et al. (2024) used machine learning models to predict the phase and crystal structure of high-entropy alloys (HEAs) [27]. Moreover, these models also facilitate the rapid screening and characterization of materials, enabling faster development cycles and improved material performance in commercial applications [28,29]. Meantime, in the field of industrial temperature prediction, Liu et al. (2024) integrated deep learning with physical models to enhance the prediction accuracy of smelting temperatures, which significantly reduced the prediction errors compared with traditional models [30]. Yang et al. (2024) proposed a graph neural network-based temperature field prediction method for steel rolling reheating furnaces, which solves the problem of irregular data containing spatial location information, which leads to inaccurate temperature field predictions [31]. Ji et al. (2024) proposed a prediction model using a hybrid of convolutional neural networks, bi-directional long short-term memory networks, and the honey badger algorithm for accurate prediction of furnace temperature during combustion in a circulating fluidized bed boiler [32]. These advances highlight the importance of combining advanced machine learning techniques with traditional process control methods to learn complex nonlinear relationships from large amounts of original data [33–35]. It has significantly improved the accuracy of temperature predictions. However, the existing literature features a significant scarcity of studies focused on applying machine learning techniques for the precise prediction of matte temperature and slag temperature in smelting furnaces. Conducting relevant research can not only address the existing research gap in this area but also provide robust technical support for optimizing and controlling the smelting process.

The objective of this work was to develop a novel hybrid nonlinear analysis algorithm for temperature control prediction in the bath smelting process. However, due to the highly nonlinear and dynamic nature of smelting processes, traditional temperature control methods often struggle to achieve accurate prediction and control. To address this issue, the proposed approach combines the strengths of the random forest–least squares support vector machine (RF-LSSVM) and random forest–relevance vector machine (RF-RVM) models, leveraging their complementary capabilities to handle the nonlinearity and complexity of the smelting process [36–38]. The integration of advanced feature extraction techniques and ensemble learning methods facilitates accurate prediction of temperature variations under the highly nonlinear and dynamic conditions characteristic of the smelting process. The innovative aspect of this work lies in the integration of multiple machine learning techniques, which not only creates a robust and adaptive predictive model but also

effectively addresses the uncertainties and complexities inherent in the smelting process. Through rigorous experimental validation, the hybrid model demonstrated substantial performance improvements in temperature prediction. This not only enhances the accuracy of temperature control in smelting operations but also offers innovative approaches and methodologies for broader industrial process optimization. The significance of this research is manifold: It not only enhances the accuracy of temperature control in smelting operations but also contributes to the broader field of industrial process optimization by demonstrating the efficacy of hybrid machine learning approaches.

The remainder of this article is structured as follows. Section 2 details the data sources and the proposed hybrid model, including the mathematical formulations and optimization strategies. Section 3 presents the results of the temperature prediction and control experiments, comparing the performance of the proposed hybrid model with traditional methods. Finally, Section 4 discusses the conclusions and the implications of the findings for industrial practice and future research directions.

2. Data and methodology

2.1. Least squares support vector machine

The LSSVM is an improvement over the standard support vector machine (SVM), characterized by its simplicity, ease of operation, fast learning speed, and ease of implementation. The principle of the LSSVM is illustrated in Figure 1. The linear regression function of the LSSVM is shown as follows:

$$y(x) = w \cdot \varphi(x) + b, \tag{1}$$

where w refers to the weight vector, $\varphi(x)$ refers to the mapping function, and b refers to the bias vector. According to the principle of structural risk minimization, the optimization problem of the LSSVM can be expressed as

$$\min_{w,b,e} J(w,e) = \frac{1}{2}w^T w + \frac{1}{2}\gamma \sum_{k=1}^{N} e_k^2, \tag{2}$$

$$y_k = w^T \varphi(x_k) + b + e_k, \tag{3}$$

where $k = 1,2,\cdots,N$, γ refers to the penalty coefficient, e_k refers to the fitting error, and b refers to the threshold. To solve this problem, the Lagrange function is constructed, and Lagrange multipliers α are introduced, such that $\alpha \geq 0$, and the optimization problem is expressed as

$$L(w,b,e,\alpha) = J(w,e) - \sum_{k=1}^{N} \alpha_k [w^T \varphi(x_k) + b + e_k - y_k]. \tag{4}$$

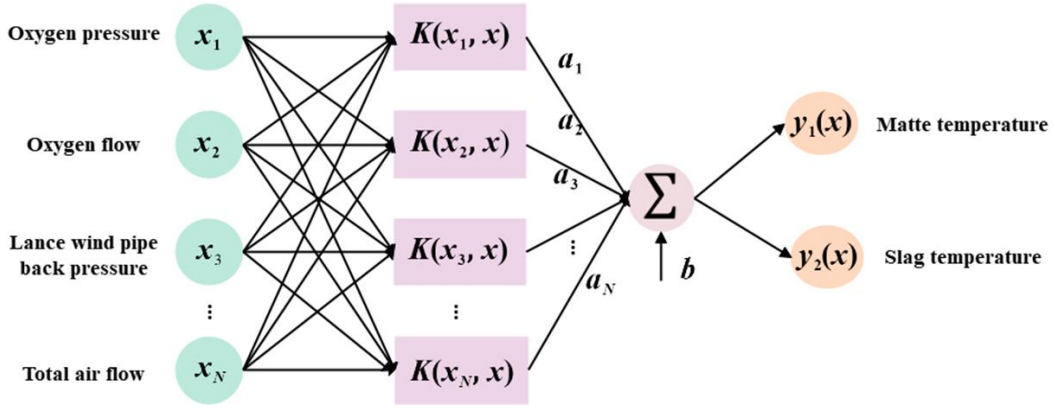

Figure 1. Schematic framework diagram of the LSSVM method.

Taking partial derivatives of the above equation,

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \quad \Rightarrow \quad w = \sum_{k=1}^{N} \alpha_k \varphi(x_k) \\ \frac{\partial L}{\partial b} = 0 \quad \Rightarrow \quad \sum_{k=1}^{N} \alpha_k = 0 \\ \frac{\partial L}{\partial e_k} = 0 \quad \Rightarrow \quad \alpha_k = \gamma e_k \\ \frac{\partial L}{\partial \alpha} = 0 \quad \Rightarrow \quad w^T \varphi(x_k) + b + e_k - y_k = 0 \end{cases}, \tag{5}$$

where $k = 1,2,\cdots,N$, thus eliminating w and e_k, and introducing the kernel function.

$$K(x_m, x_n) = \varphi(x_m)^T \varphi(x_n) \tag{6}$$

where $m, n = 1,2,\cdots,N$. Furthermore, the following matrix equation is obtained:

$$\begin{bmatrix} 0 & 1^T \\ 1 & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}, \tag{7}$$

where $1^{\mathrm{T}} = [1,1,\cdots,1]$ and $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_N]^{\mathrm{T}}$. The kernel function used in this paper is the radial basis function (RBF).

$$K(x, x_k) = \exp[-(x - x_k)^2 / 2\sigma^2]. \tag{8}$$

where σ is the kernel function's width. Finally, the prediction model of the LSSVM is obtained as follows:

$$y(x) = \sum_{k=1}^{N} \alpha_k K(x, x_k) + b. \tag{9}$$

2.2. Relevance vector machine

The relevance vector machine (RVM) shares similarities with the SVM in that it converts a linearly inseparable problem in a low-dimensional space into a linear separable problem in a high-dimensional space using a kernel function. The main difference between the RVM and the SVM is that RVM transforms hard classification into a probabilistic classification, making the

classification function maximize the likelihood function value for the training set. The principle of the RVM is illustrated in Figure 2. In RVM classification, the Laplace method is chosen for approximation, integrating to obtain the posterior probability $p(w|t,\alpha)$ of the weights and the marginal likelihood function $p(t|\alpha)$.

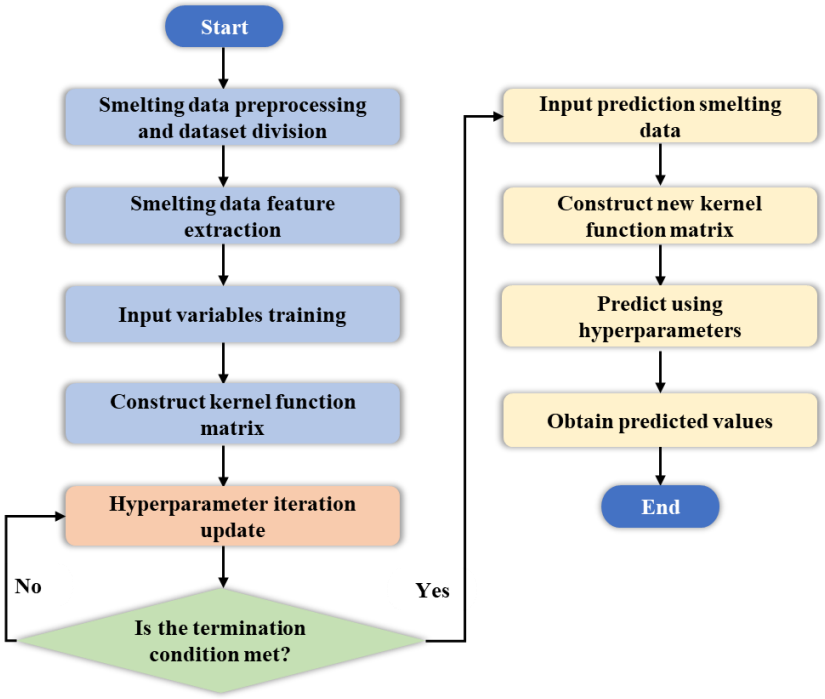

Figure 2. Schematic framework diagram of the RVM method.

Specifically, the RVM is a probabilistic model based on Bayesian principles that can learn data features, defining the prior probability influenced by the hyperparameters α over each weight ω. If the training dataset is $\{x_n, t_n | n = 1, 2, \cdots, N\}$, where x_n and t_n are the input and output values, respectively, assuming that t_n is independently distributed, the function relationship is given by

$$t_n = y(x_n; \omega) + \xi_n, \tag{10}$$

$$y(x_n; \omega) = \sum_{n=1}^{N} \omega_n K(x, x_n) + \omega_0, \tag{11}$$

where $\omega = \{\omega_n\}_{n=0}^{N}$ represents different weight values; $y(x_n; \omega)$ is a nonlinear function; $K(x, x_n)$ is the kernel function $x = (x_1, x_2, \cdots, x_N)$, where x represents a sample and x_1 is one of its features; and ξ_n is additional Gaussian noise satisfying $\xi_n \sim N(0, \sigma^2)$. The Gaussian kernel function is introduced as follows:

$$K(\|y - y_c\|) = \exp\left\{-\frac{\|y - y_c\|^2}{2\sigma^2}\right\}, \tag{12}$$

where $y_c$represents the kernel function center and σ represents the Gaussian kernel width. Assuming that t_n is independently distributed, the likelihood function is

$$p(t|\omega, \sigma^2) = (2\pi\sigma^2) - \frac{N}{2} \exp\left(-\frac{1}{2\sigma^2}\|t - \Phi\omega\|^2\right), \tag{13}$$

where $t = (t_1, t_2, \cdots, t_N)^{\mathrm{T}}$, $\omega = [\omega_0, \omega_1, \cdots, \omega_N]^{\mathrm{T}}$, and Φ is an $N \times (N+1)$ matrix.

Assuming that ω_n follows a Gaussian conditional probability distribution with a mean of 0 and the variance ω_n.

$$p(\omega|\alpha) = \prod_{n=0}^{N} N(\omega_n|0, \alpha_n^{-1}), \tag{14}$$

where α is the prior hyperparameter of the weight ω. Assuming that the hyperparameters α and the noise parameter σ^2 follow a Gamma prior probability distribution.

$$\begin{cases} P(\alpha_n) = \prod_{n=0}^{N} \mathrm{Gamma}(a, b) \\ P(\sigma^2) = \mathrm{Gamma}(c, d) \\ \mathrm{Gamma}(a, b) = \Gamma(a)^{-1} b^a \alpha^{a-1} e^{-ba} \cdot \\ \Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t}\, \mathrm{d}t \end{cases} \tag{15}$$

For more uniform hyperparameters, the common parameters are set as $a = b = c = d = 0$. Thus, the probability distribution of ω is

$$\begin{aligned} p(\omega|t, \alpha, \sigma^2) &= \frac{P\left(t|\omega, \sigma^2\right) P(\omega|\alpha)}{P\left(t|\alpha, \sigma^2\right)}, \\ =(2\pi)^{-(N+1)/2} |\Sigma|^{-1/2} &\cdot \exp\left\{-\frac{1}{2}(\omega - \mu)^{\mathrm{T}} \Sigma^{-1} (\omega - \mu)\right\}, \end{aligned} \tag{16}$$

$$\begin{cases} \Sigma = (\sigma^2 \Phi^T \Phi + A)^{-1} \\ \mu = \sigma^{-2} \Sigma \Phi^T t \end{cases}, \tag{17}$$

where Σ represents the variance, μ represents the mean, and $\mathrm{A} = \mathrm{diag}(\alpha_0, \alpha_1, \cdots, \alpha_N)$ is a diagonal matrix. Assuming the test sample y^*, the predicted value y^* is distributed as follows:

$$\begin{cases} p(t^*|t, \alpha_{\mathrm{MP}}, \sigma_{\mathrm{MP}}^2) = \int P(t^*|\omega, \sigma_{\mathrm{MP}}^2) P(\omega|t, \alpha_{\mathrm{MP}}, \alpha_{\mathrm{MP}}) \mathrm{d}\,\omega \\ p(t^*|t, \alpha_{\mathrm{MP}}, \sigma_{\mathrm{MP}}^2) = N(t^*|y^*, \sigma_*^2) \end{cases}, \tag{18}$$

where the variance $\sigma_*^2 = \sigma_{\mathrm{MP}}^2 + \varphi^{\mathrm{T}}(y^*) \sum \varphi(y^*)$ and the predicted value y^* of the test sample t^* is distributed as $f^* = (y^*; \mu)$.

The kernel width and regularization parameters performed by the two algorithms mentioned above are optimized to enhance the predictive performance of the LSSVM and RVM algorithms for matte and slag temperatures. During evolution, we recorded the fitness $f(x_i(t))$ of each individual $x_i(t)$ ($i = 1, \cdots, N, t$ current algebra) along with the corresponding parameter combinations $(\sigma_i(t), \lambda_i(t))$. When a subpopulation $P_s(t)$ reaches a local optimum, assuming that the current optimal individual is $x_{best}^s(t)$, the individual x_{ref} and its parameter combinations $(\sigma_{ref}, \lambda_{ref})$, which exhibit similar fitness to $x_{best}^s(t)$ and in a different parameter space, are identified by searching through previously recorded historical data.

In this work, a differential evolutionary algorithm is implemented using numerical analysis techniques. At the conclusion of the algorithm, the parameter values corresponding to the optimal individuals are recorded, and the algorithm is executed multiple times to obtain stable range values.

If the matte temperature (Y_1) is predicted during one of the runs, σ stabilizes between 0.8 and 1.2, while λ stabilizes between 0.4 and 0.6. Under these conditions, the LSSVM and RVM models demonstrate reduced prediction errors for the matte and slag temperatures, resulting in a better fit.

2.3. Random forest

The random forest (RF) can be used to evaluate the importance of features, that is, the contribution of each feature to the predictive ability of the model. The principle of the RVM is illustrated in Figure 3. One of the methods of calculating feature importance is to observe the change in model performance through disrupting the value of the features. Specifically, for each feature A_j, the average value of the difference between the prediction errors of all out-of-bag samples before and after disrupting the feature's value is calculated. The feature importance score can be expressed as follows:

$$Importance(A_j) = \frac{1}{N}\sum_{i=1}^{N} [Error(x_i^{OOB}) - Error(x_i^{\text{OOB,permuted}})], \tag{19}$$

where $\text{Error}(x_i^{\text{OOB}})$ refers to the original prediction error of the i^{th} out-of-bag sample $\text{Error}(x_i^{\text{OOB,permuted}})$. RF effectively improves the performance of the model by constructing multiple decision trees based on random feature subsets and integrating their prediction results. Its core idea is to use the diversity of ensemble learning to reduce variance and bias and enhance the generalization ability and anti-overfitting ability of the model. Through out-of-bag sample evaluation and feature importance evaluation, random forest also provides a powerful tool for model performance evaluation and feature selection.

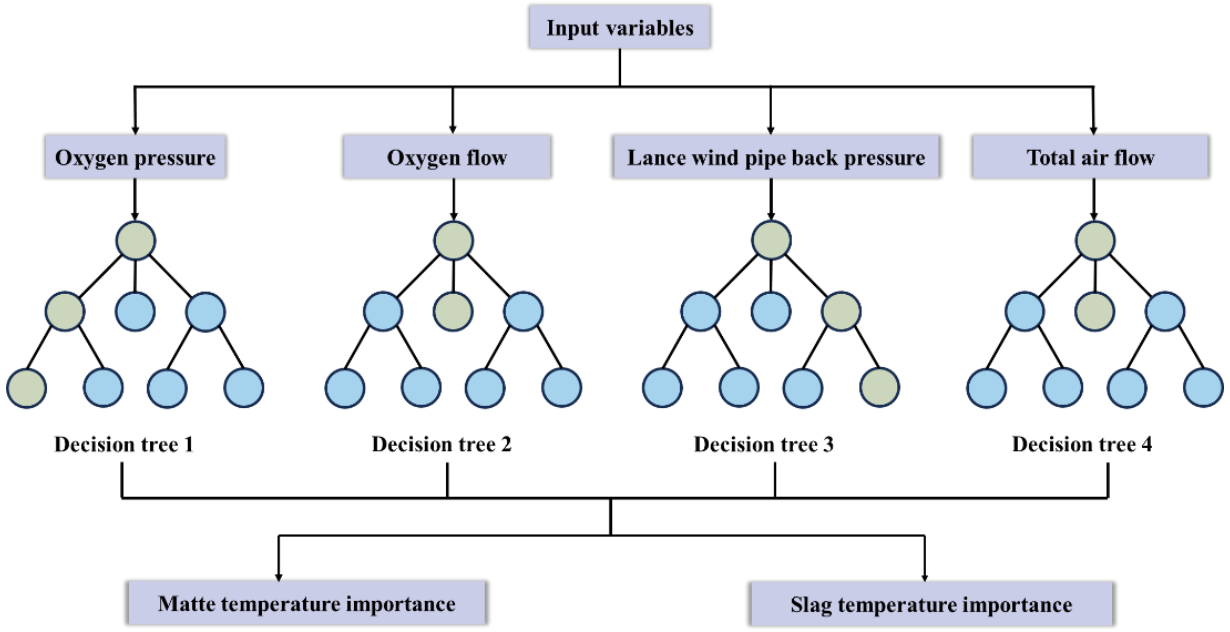

Figure 3. Schematic framework diagram of the RF method.

2.4. Data sources

A certain company uses an oxygen-enriched top-blown furnace for nickel metal smelting.

Depending on the ore grade, the content of substances in the matte varies. The matte produced by the oxygen-enriched top-blown furnace mainly contains nickel, copper, iron, and sulfur, while the slag mainly contains magnesium oxide and calcium oxide. The furnace body is a vertical cylinder constructed from a high-strength steel plate, lined with refractory material, and divided into a flue gas area and a smelting pool area. The bottom features outlets for slag and metal discharge. A feed port and a lance port are included in a water-cooled membrane-type wall cover on the top of the furnace. The core component of the lance assembly is made up of multi-layer concentric tubes that have a cooling system. Vertical adjustment of this assembly is necessary for optimal positioning. The flue gas exhaust system is equipped with a flue gas outlet and purification equipment to minimize environmental pollution. The fundamental working principle of the oxygen-enriched top-blowing furnace involves the use of a lance that is inserted vertically into the molten pool to accurately inject oxygen-enriched air and fuel. During this process, a strong air stream agitates the molten pool, leading to a series of complex physicochemical changes, including smelting, sulfurization, oxidation, and reduction, all occurring in a high-temperature environment. These transformations enable the oxygen-enriched top-blowing furnace to efficiently carry out the smelting and refining of metals.

Due to the high temperature and complex equipment environment during the smelting process, some data were collected during the production process of this oxygen-enriched top-blown furnace, such as the feeding amount (X_1), oxygen pressure (X_2), oxygen flow (X_3), pipeline pressure (X_4), lance windpipe back pressure (X_5), lance windpipe flow (X_6), total air flow (X_7), oxygen concentration (X_8), exhaust gas residual oxygen concentration (X_9), matte temperature (Y_1), slag temperature (Y_2), and slag (liquid) level height (Y_3). Among them, feeding amount (X_1), oxygen pressure (X_2), oxygen flow (X_3), pipeline pressure (X_4), lance windpipe back pressure (X_5), lance windpipe flow (X_6), total air flow (X_7), oxygen concentration (X_8), and exhaust gas residual oxygen concentration (X_9) are the input parameters (independent variables). Matte temperature (Y_1) and slag temperature (Y_2) are the output parameters (dependent variables). The data were collected every hour, and a total of 868 sets of data were collected from 1 July 2023 to 5 August 2023, as shown in Table 1.

Table 1. Statistics of the processing parameters for the oxygen-enriched top-blown furnace system.

No. (unit)	Parameter	Mechanistic descriptions
X_1 (t/h)	Feeding amount	The quantity of raw material introduced into the furnace or other smelting equipment per unit of time
X_2 (kPa)	Oxygen pressure	The pressure of oxygen supplied to a smelting furnace or other smelting equipment
X_3 (Nm^3/h)	Oxygen flow	The volume of oxygen delivered to the furnace or other smelting equipment per unit of time
X_4 (kPa)	Pipeline pressure	Pressure of a fluid in a conveying pipe
X_5 (kPa)	Lance windpipe back pressure	The resistance pressure of the gas within the gun duct during the flow process
X_6 (Nm^3/h)	Lance windpipe flow	The volume of gas passing through the gun duct per unit of time
X_7 (Nm^3/h)	Total air flow	The total volume of gas passing through the blast system of a melting furnace per unit of time
X_8 (%)	Oxygen concentration	Percentage of oxygen by volume in the gas mixture
X_9 (%)	Exhaust gas residual oxygen concentration	The percentage by volume of oxygen remaining in the exhaust gases emitted from the smelting process
Y_1 (℃)	Matte temperature	Real-time temperature of the matte during melting
Y_2 (℃)	Slag temperature	Real-time temperature of slag during the melting process

2.5. Data preprocessing

Due to the presence of a small amount (less than 1%) of missing values in the production data collected from the company, it was necessary to perform data cleaning on the raw data, such as deleting rows or columns with missing values. Since the data collected for each parameter are time series data, linear interpolation was used to fill in missing values. Linear interpolation is a simple and commonly used interpolation technique. Suppose there are two known data points, (x_1, y_1) and (x_2, y_2), and we want to estimate the value yy corresponding to a point xx between them. Linear interpolation assumes that the change between the two known points is linear, i.e., (x_1, y_1) and (x_2, y_2) can be connected by a straight line. The formula for linear interpolation to calculate the missing value y is

$$y = y_1 + \frac{(x-x_1)}{(x_2-x_1)} \times (y_2 - y_1). \tag{20}$$

Simultaneously, to enhance both the convergence speed and accuracy of the model, data normalization is essential. Each factor takes N=700 representative test data. Since all the data are

deterministic, min-max normalization represents a linear transformation of the raw data, ensuring that the results are scaled within the [0, 1] range. Therefore, the min-max normalization method is used to process the data as follows:

$$x_i = \frac{X_i - X_i^{\min}}{X_i^{\max} - X_i^{\min}}, \tag{21}$$

$$y_j = \frac{Y_j - Y_j^{\min}}{Y_j^{\max} - Y_j^{\min}}, \tag{22}$$

where $x_i(i = 1,2,\cdots,9)$ is the transformed input parameter, $y_j(j = 1,2,3)$ is the transformed output parameter, $y_j(j = 1,2,3)$ is the original input parameter mentioned earlier, $y_j(j = 1,2,3)$ is the original output parameter mentioned earlier, $X_i^{\max}$ is the maximum value of the i^{th} influencing factor in the test data, $X_i^{\min}$ is the minimum value of the j^{th} influencing factor in the test data, $Y_j^{\max}$ is the maximum value of the j^{th} dependent variable in the test data, and $Y_j^{\min}$ is the minimum value of the j^{th} dependent variable in the test data.

Specifically, the box plots of the normalized values of matte temperature and slag temperature and their influencing factors are shown in Figure 4 andFigure **5**. It can be seen that for matte temperature (Y_1) and slag temperature (Y_2), there are data points outside the left and right boundary points (considered as outliers). The median is close to the center of the box, indicating that the data are relatively uniformly distributed, but the shape of the data distribution varies for each column. The normalized values of matte temperature and slag temperature show wider boxes, indicating greater variability in these parameters. As indicated in Figure 5, it can be seen that most of the independent variables have some outliers, represented by discrete points in the figure. Taking oxygen pressure (X_2) and lance windpipe flow (X_6) as examples, the median is close to the center of the box, indicating that these data are relatively uniformly distributed.

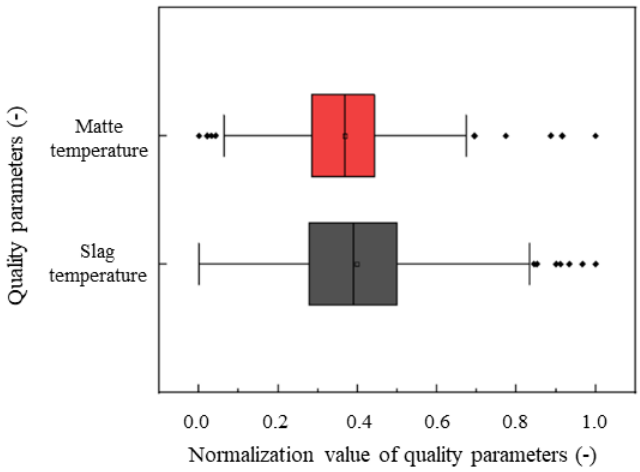

Figure 4. Box diagram of the normalized value of quality parameters of oxygen-rich top-blown smelting.

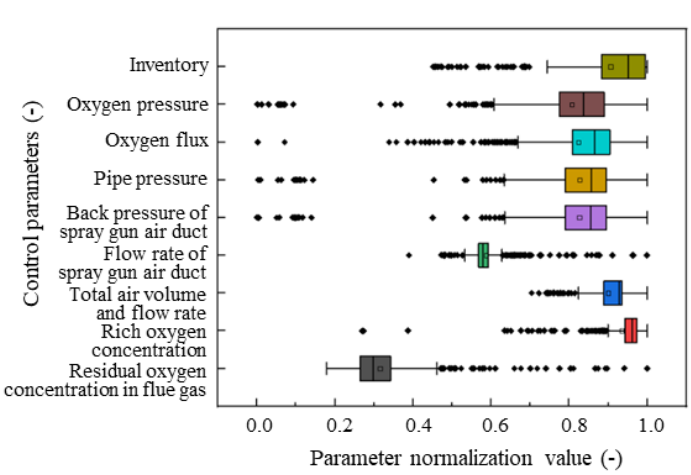

Figure 5. Box diagram of normalized control parameters for oxygen-rich top-blown smelting.

The evaluation of the prediction results uses the mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE), where the MAE can well reflect the actual situation of prediction error, MAPE can well reflect the relative error of a prediction by reflecting the percentage average of the error between the predicted, and actual values. and the RMSE is used to measure the deviation between the predicted values and true values. For the j^{th} component of smelting quality, the calculation formulae are as follows:

$$\begin{cases} \text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i(k) - y_i(k)|^2 \\ \text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i(k) - \hat{y}_i(k)\right)^2} \\ \text{MAPE} = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i(k) - y_i(k)}{y_i}\right| \\ \text{R}^2 = 1 - \frac{\sum_{i=1}^{N}|\hat{y}_i(k) - y_i(k)|^2}{\sum_{i=1}^{N}|y_i(k) - \bar{y}_i(k)|^2} \end{cases}, \tag{23}$$

where $y_j(k)$ represents the true value of the i^{th} component of smelting quality in the k^{th} experiment and $\hat{y}_j(k)$ represents the predicted value of the i^{th} component of smelting quality in the k^{th} experiment.

2.6. Hybrid model mechanism

Figure 6 illustrates the framework of temperature control prediction of the bath smelting process. The proposed RF-LSSVM and RF-RVM models integrate the strengths of RF for feature selection and the LSSVM/RVM for nonlinear regression. RF employs an ensemble of decision trees to evaluate features' importance. Key parameters such as the feeding amount (X_1) and oxygen pressure (X_2) are identified as the dominant factors, reducing redundant variables and enhancing the model's interpretability.

The LSSVM utilizes the RBF kernel to project data into a high-dimensional space, where linear separation becomes feasible. The penalty coefficient balances model complexity and fitting error. The RVM adopts a Bayesian framework to automatically select relevant vectors, ensuring sparsity and robustness against overfitting. The Gaussian kernel width is optimized via evidence maximization. In terms of robustness against noise, the ensemble nature of RF mitigates outliers in

industrial data, while the structural risk minimization of the LSSVM and the probabilistic output of the RVM enhance adaptability to dynamic smelting conditions.

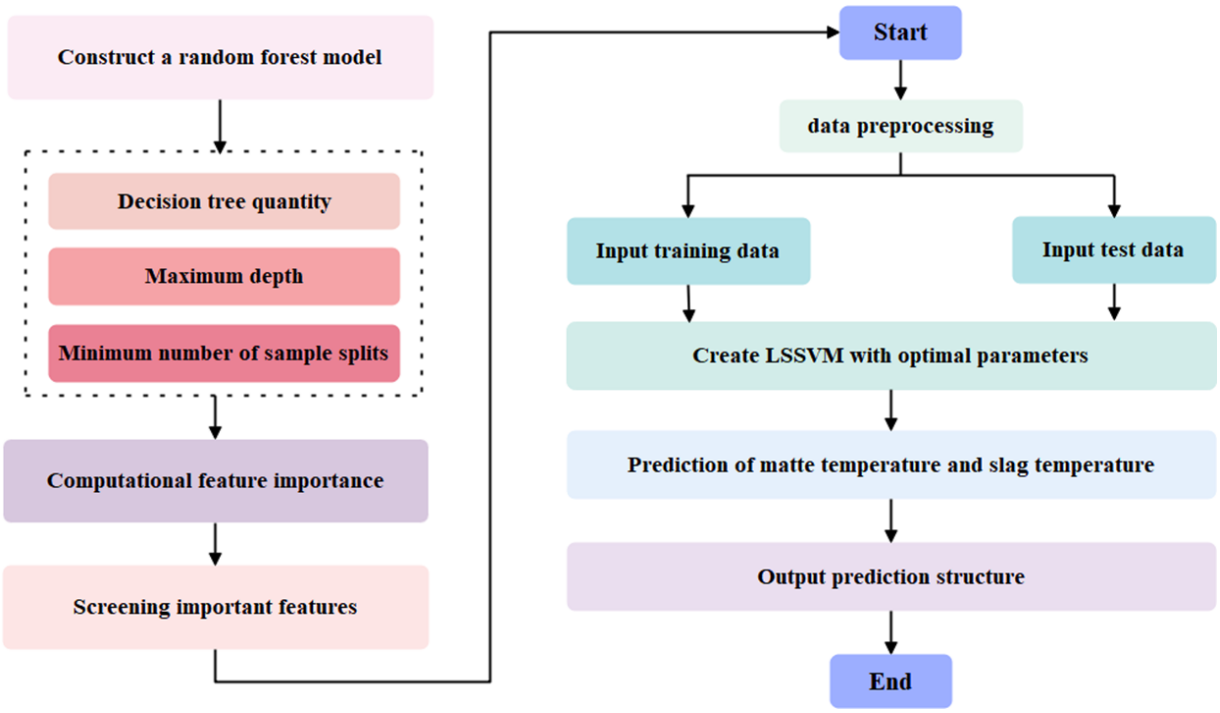

Figure 6. Predictive framework for the matte temperature and slag temperature.

3. Results and discussion

3.1. Nonlinear correlation analysis

The maximal information coefficient (MIC) is a statistical method used to quantify the strength of the relationship between two variables. It is part of a series of statistics known as maximal information-based nonparametric exploration, designed to capture various types of relationships, including linear, nonlinear, and complex patterns. The calculation of the MIC does not have a simple closed formula, but it is based on a relatively complex algorithm. The general steps of this algorithm can be summarized as follows.

Step 1: Create a grid for the two variables. For example, given two variables X and Y, first create a grid in the joint space.

Step 2: For each possible grid division, calculate the mutual information $I(X:Y)$ of X and Y. The mutual information formula is as follows:

$$I(X:Y) = \sum_{x\in X, y\in Y} P(x,y)\log\left(\frac{P(x,y)}{P(x)P(y)}\right) \tag{23}$$

where $P(x,y)$ is the joint probability distribution of X and Y, and $P(x)$ and $P(y)$ are the marginal probability distributions of X and Y, respectively.

Step 3: For each grid division, calculate the normalized mutual information $I^*(X:Y)$, which is the ratio of mutual information to the maximum possible mutual information. This ratio is used to

standardize the grid size.

Step 4: MIC is defined as the maximum value of $I^*(X:Y)$ among all grid divisions, as follows:

$$\mathrm{MIC}(X:Y) = \max_{grid\ divisions} I^*(X:Y). \tag{24}$$

The key to this process is to try many different grid divisions in the space, calculate the normalized mutual information for each division method, and then select the maximum value as the MIC value. In fact, MIC values between 0.90 and 1.00 indicate extremely high correlation, values between 0.70 and 0.90 indicate high correlation, values between 0.40 and 0.70 indicate moderate correlation, values between 0.20 and 0.40 indicate low correlation, values between 0.10 and 0.20 indicate very low correlation, and values less than 0.10 indicate no correlation. Therefore, the MIC is introduced in this paper to measure the nonlinear strength between factors influencing smelting quality, playing an important role in systematically revealing the driving factors of smelting quality control.

For the two dependent variables (matte temperature (Y_1) and slag temperature (Y_2)) and their nine different influencing factors (feeding amount (X_1), oxygen pressure (X_2), oxygen flow (X_3), pipeline pressure (X_4), lance windpipe back pressure (X_5), lance windpipe flow (X_6), total air flow (X_7), oxygen concentration (X_8), and exhaust gas residual oxygen concentration (X_9)), 868 sets of collected data were analyzed using a mathematical software program code to run the MIC statistical method. The software's source code comes from Albanese et al., and The MIC calculation results are shown in Figure 7. The results show that pipeline pressure (X_4) and lance windpipe back pressure (X_5) exhibit an extremely high correlation, with an MIC of 0.9150. This is because gas is blown through the pipeline at a certain pressure, and when the airflow is obstructed, it creates back pressure on the lance windpipe, showing a close relationship. Additionally, oxygen pressure (X_2) shows high correlations with oxygen flow (X_3), pipeline pressure (X_4), and lance windpipe back pressure (X_5), with MIC values of 0.7487, 0.7530, and 0.7545, respectively. Oxygen flow (X_3) shows high correlations with pipeline pressure (X_4) and lance windpipe back pressure (X_5), with MIC values of 0.7368 and 0.7385, respectively. Other influencing factors show moderate correlations with each other. This indicates that by fully utilizing the fluctuation characteristics of data in the oxygen-enriched top-blown smelting process, the relationships between the factors influencing smelting quality can be better analyzed. The model-independent nature of the maximal information coefficient is well-suited for exploring the relationships between complex variables such as fluctuations in influencing factors of smelting quality. When formulating smelting quality control measures, the dynamic relationships between different influencing factors should be considered comprehensively, and excessive intervention in any single influencing factor should be avoided. It is noteworthy that the MIC values between the nine different influencing factors and the two dependent variables, matte temperature (Y_1) and slag temperature (Y_2), are distributed between 0.30 and 0.70, showing moderate and low correlation. This indicates that matte temperature (Y_1) and slag temperature (Y_2) are not influenced by any single factor, and their main control factors need to be further explored.

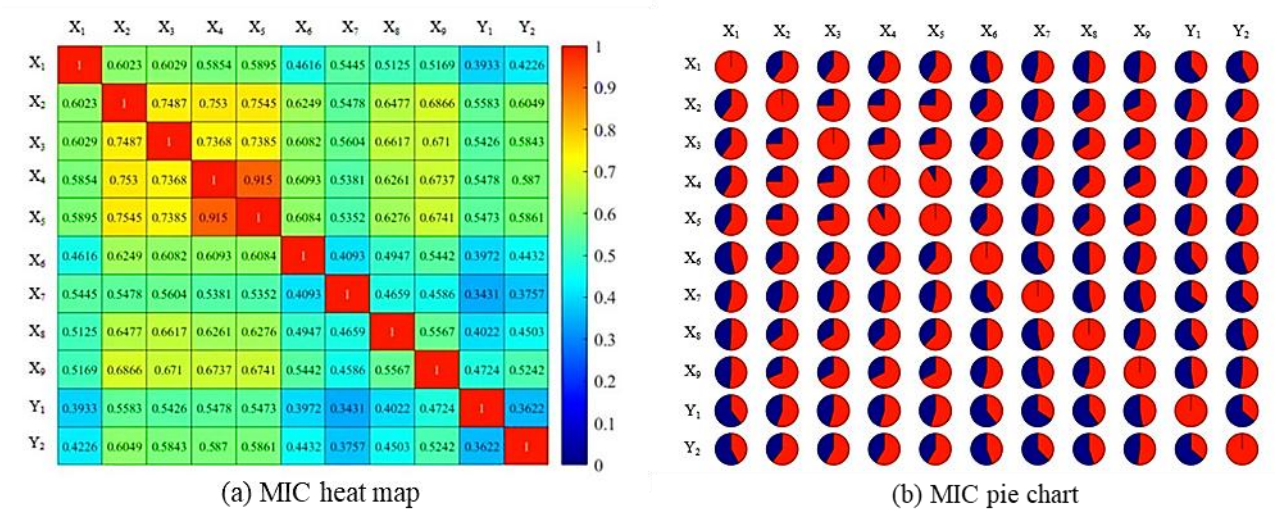

Figure 7. Nonlinear analysis results of factors affecting the temperature using MIC.

3.2. Identification of main control factors

To further investigate the factors influencing smelting quality and enhance the interpretability of the model, the random forest method, a widely adopted ML technique, is employed for nonlinear modeling. By integrating multiple decision trees, this approach enables data classification, correlation testing, prediction generation, and result interpretation. In this work, the dependent variable smelting quality is analyzed using RF to identify and understand the independent variables that most significantly affect the dependent variables, matte temperature (Y_1) and slag temperature (Y_2). By offering insights into the interrelationships among variables within the dataset, identifying the dominant factors influencing model predictions enhances both data comprehension and an understanding of the underlying smelting-related issues.

Hence, we used the RF algorithm to analyze the importance of factors affecting smelting quality. The output results of the algorithm are shown in Table 2. In this table, the top five important indicators affecting matte temperature (Y_1) are 1.0361, 0.8824, 0.8007, 0.7917, and 0.7648. The top five important indicators affecting slag temperature (Y_2) are 0.9025, 0.7815, 0.6559, 0.6553, and 0.5873. Therefore, the top five main control factors for matte temperature (Y_1) are total air flow (X_7), oxygen flow (X_3), lance windpipe back pressure (X_5), feeding amount (X_1), and oxygen pressure (X_2). The top five main control factors for slag temperature (Y_2) are pipeline pressure (X_4), total air flow (X_7), oxygen flow (X_3), oxygen pressure (X_2), and lance windpipe back pressure (X_5). The main control factors for both matte temperature (Y_1) and slag temperature (Y_2) include oxygen pressure (X_2), oxygen flow (X_3), lance windpipe back pressure (X_5), and total air flow (X_7). This suggests that the impact of the lance on the molten pool is a critical factor influencing smelting quality, particularly as the stirring effect induced by blowing oxygen plays a key role in enhancing the uniformity of the temperature field within the molten pool, which, in turn, affects smelting quality. Further studies have demonstrated that utilizing MIC to analyze the nonlinear correlations among the factors provides additional support for the primary control factors in smelting quality control identified through the RF method.

Table 2. The importance of factors affecting the temperature.

Factor	Y_1	Y_2
X_1	0.7917	0.2730
X_2	0.7648	0.6553
X_3	0.8824	0.6559
X_4	0.7254	0.9025
X_5	0.8007	0.5873
X_6	0.1244	0.2285
X_7	1.0361	0.7815
X_8	0.6037	0.5434
X_9	0.4346	0.0073

3.3. Effect of different prediction methods on prediction accuracy

Building on the identification of the key influencing factors, this section compares the performance of various prediction models in forecasting smelting quality. Using the two smelting quality indicators, matte temperature (Y_1) and slag temperature (Y_2), as well as the nine influencing factors affecting smelting quality as the sample database, N_1 sets of smelting data are used as the training set, and the remaining N_2 sets are used as the test set. Quality control prediction models for smelting are established using the LSSVM, RVM, XGBoost, multiple linear regression (MLR), back propagation (BP) the neural network (BP-NN), the kernel extreme learning machine (KELM), and long short-term memory (LSTM). The three main calculation parameters of the LSSVM are kernel width sig2 = 500 and regularization parameter gamma = 5, and the chosen kernel function is the RBF kernel function. An integrated approach combining 10-fold cross-validation and differential evolution (DE) algorithms is used to optimize the radial basis function (RBF) kernel parameter σ. Through an initial analysis of the input demand space, the search range of σ is determined to be [X, Y], which ensures coverage of ±3 standard deviations of the normalized feature distance. The RVM also uses RBF as the kernel function (i.e., Gaussian kernel function) with a kernel width of =0.1. XGBoost is based on a linear kernel, with an L1 regularization parameter of 0.1, an L2 regularization parameter of 0.1, 1000 iterations, and a learning rate of 0.01. The main parameter of MLR is the regression coefficient (also known as the weight), which is estimated by fitting the training data. The BP neural network is set to a three-layer network structure with a maximum of 1000 iterations, a learning rate of 0.01, a training error of 0.0001, a momentum factor of 0.01, a minimum performance gradient of 10^{-6}, and a maximum failure count of 6. These five methods were used to model and predict 868 consecutive actual production data.

Furthermore, Figure 8 illustrates the comparative results of the matte temperature data and the slag temperature data predicted by seven different models. The black curves in Figure 8(a) represent the true values of the matte temperature, and the other colored curves represent the predicted values of the matte temperature. The black curves in Figure 8(b) represent the true values of the slag temperature, and the other colored curves represent the predicted values of the slag temperature. According to Figure 8(a), each prediction method has some error in the predicted values of matte temperature. However, the MLP method and BP-NN method have a larger error in the predicted values of the matte temperature, and the LSSVM method has a smaller error in the predicted values of the matte temperature. It can be concluded that the LSSVM model is more accurate in predicting the matte temperature and better reflects the future trend of the complex nonlinear matte temperature

data. According to Figure 8(b), each prediction method has some error in the predicted values of slag temperature. However, the MLP method and the BP-NN method have a larger error in the predicted values of the slag temperature, and the LSSVM method has a smaller error in the predicted values of the slag temperature. It can be concluded that the LSSVM model is more accurate in predicting the slag temperature and better reflects the future trend of the complex nonlinear slag temperature data.

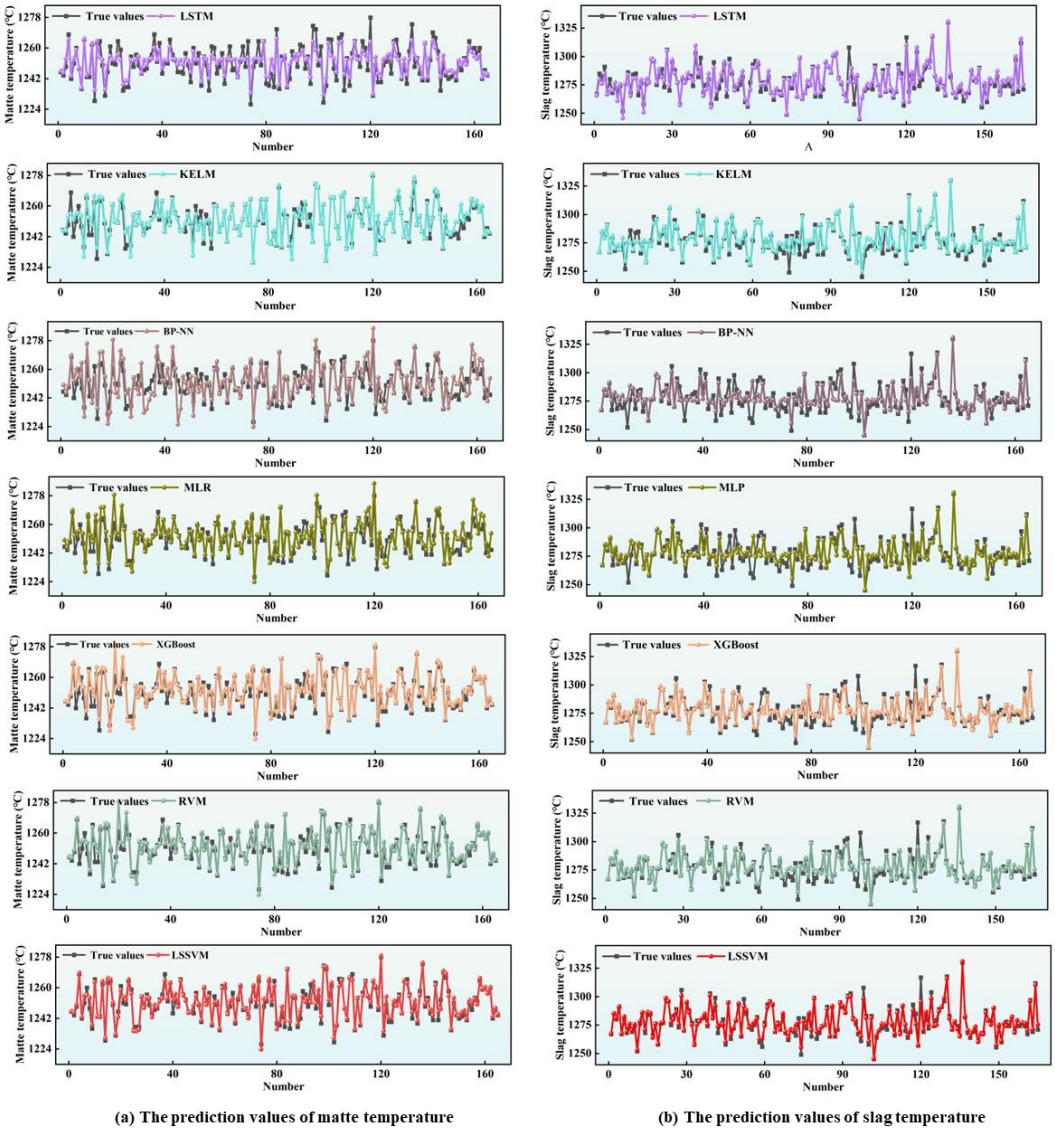

Figure 8. Comparison of the predicted values of matte temperature (a) and slag temperature (b) from different prediction methods.

It is worth noting that this study utilizes four evaluation metrics, specifically MAE, RMSE, MAPE, and R^2, to comprehensively assess the prediction performance of seven methods regarding matte temperature and slag temperature. This approach establishes a more objective benchmark for comparing the predictive capabilities of these methods in the context of matte and slag temperature forecasting. Table 3 displays the results of the evaluation indexes for each prediction method. In this table, the best prediction performance of the seven prediction methods on matte temperature data is that of the LSSVM method, with R^2 reaching 0.93. The best prediction performance of the seven prediction methods for slag temperature data is that of the LSSVM method, with R^2 reaching 0.94. It

is concluded that the four evaluation indexes are more intuitive to show that the LSSVM method has a better prediction ability for matte temperature and slag temperature.

Table 3. Evaluation of the prediction results of temperature for different prediction methods.

Smelting quality	Indicators	LSSVM	RVM	XGBoost	MLR	BP-NN	KELM	LSTM
Y_1	MAE	7.87	8.32	11.9406	8.23	23.64	11.50	8.52
	RMSE	9.95	10.21	9.56	10.31	55.92	9.54	10.45
	MAPE	15.70%	18.27%	20.49%	31.46%	43.31%	16.83%	30.98%
	R^2	0.93	0.89	0.87	0.74	0.54	0.88	0.78
Y_2	MAE	10.19	10.63	15.55	10.70	55.92	10.50	15.10
	RMSE	13.11	13.44	12.9611	13.58	59.14	13.25	12.56
	MAPE	17.45%	20.50%	25.41%	36.19%	41.99%	21.49%	25.32%
	R^2	0.94	0.91	0.87	0.75	0.68	0.91	0.88

3.4. Effect of different training–testing set proportions on prediction accuracy

In fact, the sample size of the parameters collected in the actual production process is often as large as possible. However, the quantity of data utilized for predicting the status of smelting quality control remains uncertain, necessitating a sensitivity analysis regarding the number of predicted data points based on the evaluation metrics presented in Section 2 (i.e., MAE and RMSE). When the proportion of the training set is 80%, 85%, 90%, and 95%, the sample sizes of the training set and test set are 695, 738, 781, and 825 and 173, 130, 87, and 43, respectively. This work compares and analyzes the sensitivity of the number of predicted values under different indicators of the status of smelting quality control. When the proportion of the training set sample size to the total dataset changes from 80% to 95%, the MAE and RMSE of the RVM prediction model for smelting quality do not fluctuate repeatedly but show a trend of first decreasing and then increasing, with the minimum values appearing when the training set is 80%.

For the LSSVM model presented in Table 4, the lowest MAE value of 7.66 is achieved with an 85% training set, while the lowest RMSE of 9.50 also occurs at this ratio. The highest R^2 value of 0.93 is achieved with an 80% training set. For Y_2, the lowest MAE of 10.19 and RMSE of 13.11 are observed for the 80% training set, which also yields the highest R^2 value of 0.94. As shown in Table 5, the RVM model indicates that when predicting Y_1, the lowest MAE of 8.32 is achieved with an 80% training set, while the lowest RMSE of 10.16 is achieved with an 85% training set. The highest R^2 of 0.89 is also observed with the 80% training set. For the second outcome, Y_2, the lowest MAE of 9.98 is observed with the 95% training set, and the lowest RMSE of 12.14 is also achieved for this same ratio. Conversely, the highest R^2 of 0.94 is observed with the 95% training set.

Table 4. Effect of prediction step size on the matte and slag temperature predicted by the LSSVM.

Parameter	Indicators	75%	80%	85%	90%	95%
Y_1	MAE	8.76	7.87	7.66	8.21	8.94
	RMSE	10.52	9.95	9.50	10.03	10.89
	MAPE	18.96%	15.70%	16.40%	16.00%	19.5%
	R^2	0.87	0.93	0.91	0.86	0.72
Y_2	MAE	10.36	10.19	10.75	11.38	11.87
	RMSE	13.66	13.11	13.40	13.87	14.74
	MAPE	19.00%	17.45%	18.10%	19.50%	20.50%
	R^2	0.88	0.94	0.91	0.84	0.77

Table 5. Effect of predicted step size on matte temperature and slag temperature predicted by the RVM.

Parameter	Indicators	75%	80%	85%	90%	95%
Y_1	MAE	8.82	8.32	8.45	9.09	8.77
	RMSE	11.89	10.21	10.16	11.33	11.40
	MAPE	19.20%	18.27%	18.54%	20.00%	19.00%
	R^2	0.86	0.89	0.88	0.85	0.87
Y_2	MAE	11.71	10.63	11.66	10.30	9.98
	RMSE	14.83	13.44	14.47	13.15	12.14
	MAPE	25.40%	20.50%	22.41%	18.50%	17.50%
	R^2	0.81	0.91	0.89	0.93	0.94

Cross-validation is a highly effective technique for evaluating the performance and robustness of a model. Its core principle involves dividing the original dataset into multiple training and validation sets. By repeatedly training and validating the model on different divisions of the data, we can assess its overall performance and reduce bias associated with data partitioning. One commonly used method is k-fold cross-validation, where the process is repeated k times, ensuring that each subset serves as a validation set at least once. The final evaluation metric of the model is calculated by averaging the performance metrics obtained from the k validation results. This approach enables a comprehensive assessment of model performance across different data subsets, offering a more reliable measure of robustness and generalization capability.

The evaluation process of cross-validation algorithm is as follows. Let the LSSVM model be M, and the training set at the ith validation is T_i and the validation set is V_i, $i = 1,2,\cdots,k$. The performance index of the model at the i^{th} validation is as follows:

$$MSE_i = \frac{1}{|V_i|}\sum_{x_j \in V_i}\left(y_j - \hat{y}_j\right)^2 \tag{25}$$

where V_i is the number of samples in the validation set V_i, y_j is the true value of the sample x_j, and $\hat{y}_j$ is the predicted value of the model M for the sample x_j.

$$MSE_{avg} = \frac{1}{k}\sum_{i=1}^{k} MSE_i. \tag{26}$$

3.5. Impact of main control factors on prediction accuracy

To further enhance prediction performance and the models' interpretability, hybrid models combining feature selection and machine learning were constructed. Using N_1 = 695 points of smelting data as the training set and the remaining N_2 = 173 as the test set, the RF-LSSVM and RF-RVM models were established to intelligently predict the temperature. The comparison results of the RF-LSSVM and LSSVM models in predicting the matte and slag temperatures are shown in Table 6. When predicting the matte temperature (Y_1), the MAE and RMSE of the RF-LSSVM model are lower than those of the LSSVM model. Additionally, the MAPE decreases from 18.27% to 14.00%, and the R^2 improves from 0.89 to 0.94. Similarly, when predicting slag temperature (Y_2), the RF-LSSVM model achieves an MAE of 10.47 and an RMSE of 13.31, both lower than the corresponding values of the LSSVM model. Moreover, the MAPE decreases from 20.50% to 17.00%, and the R^2 increases from 0.91 to 0.95. These results indicate that the RF-LSSVM model yields smaller prediction errors, better fitting performance, and higher prediction accuracy. The RF and LSSVM play different and complementary roles in predicting matte and slag temperatures. RF provides more precise and representative input data for the LSSVM through feature selection and preliminary prediction, which reduces the number of features to be processed by the LSSVM, reduces the complexity of the model, and also provides a general direction for training the LSSVM, which helps the LSSVM to converge to the optimal solution faster.

Additionally, the comparison results of the RF-RVM and RVM models in predicting the matte and slag temperatures are shown in Table 7. For the matte temperature (Y_1), the MAE and RMSE of the RF-RVM model are lower than the corresponding values for the RVM model, which are 8.32 and 10.21, respectively. Additionally, the MAPE decreases from 18.27% to 15.00%, and the R^2 improves from 0.89 to 0.92. For the slag temperature (Y_2), the RF-RVM model yields an MAE of 10.53 and an RMSE of 13.38, both lower than the corresponding values of the RVM model, which are 10.63 and 13.44, respectively. The MAPE decreases from 20.50% to 18.00%, and the R^2 improves from 0.91 to 0.93. These results indicate that the prediction accuracy and goodness of fit of the RF-RVM model are significantly enhanced compared with the RVM model. The reason for this is the improved clarity, readability, and technical accuracy while maintaining the original meaning. The feature selection and preprocessing of RF provide more representative and concise input data for the RVM, which reduces the complexity of RVM processing and helps the RVM find the nonlinear relationships in the data faster. The powerful nonlinear modeling capability and sparsity property of the RVM compensate for the lack of accuracy and high model complexity of RF in dealing with complex nonlinear relationships. For instance, after RF identifies the key features that affect the matte temperature, the RVM is able to more accurately represent the nonlinear mapping relationship between these features and the matte temperature and, at the same time, reduce the redundant information in the model through sparsity selection to improve the accuracy and efficiency of the model. The comparison also shows that the prediction accuracy of the RF-LSSVM model for the matte and slag temperatures is higher than that of the RF-RVM model.

Table 6. Comparison of prediction results for matte and slag temperature of the RF-LSSVM and LSSVM.

Indicators	MAE		RMSE		MAPE		R^2	
Model	LSSVM	RF-LSSVM	LSSVM	RF-LSSVM	LSSVM	RF-LSSVM	LSSVM	RF-LSSVM
Y_1	8.32	7.58	10.21	9.82	18.27%	14.00%	0.89	0.94
Y_2	10.63	10.47	13.44	13.31	20.50%	17.00%	0.91	0.95

Table 7. Comparison of predictions of the matte and slag temperature of the RF-RVM and RVM.

Indicators	MAE		RMSE		MAPE		R^2	
Model	RVM	RF-RVM	RVM	RF-RVM	LSSVM	RF-LSSVM	LSSVM	RF-LSSVM
Y_1	8.32	7.79	10.21	9.78	18.27%	15.00%	0.89	0.92
Y_2	10.63	10.53	13.44	13.38	20.50%	18.00%	0.91	0.93

When N_1 = 695 sets of smelting data were used as the training set and the remaining N_2 = 173 sets as the test set, the RF-LSSVM and RF-RVM models were further established to intelligently predict the matte temperature and slag temperature in metallurgical engineering. Table 8 shows the impact of the common main control factors on the accuracy of the RF-LSSVM and RF-RVM models for predicting matte temperature and slag temperature. From the table, it can be seen that for matte temperature (Y_1), the MAE and RMSE of the RF-LSSVM algorithm based on the four main control factors are 2.64% and 1.02% lower, respectively, compared with the original RF-LSSVM algorithm. For slag temperature (Y_2), the MAE and RMSE of the RF-LSSVM algorithm based on the four main control factors are 0% and 0.45% lower, respectively, compared with the original RF-LSSVM algorithm. For matte temperature (Y_1), the MAE and RMSE of the RF-RVM algorithm based on the four main control factors are 1.16% and 1.63% lower, respectively, compared with the original RF-RVM algorithm. For slag temperature (Y_2), the MAE and RMSE of the RF-RVM algorithm based on the four main control factors are 1.61% and 0% higher, respectively, compared with the original RF-RVM algorithm. Specifically, the prediction results of the RF-LSSVM model and the RF-RVM model based on the four main control factors for matte temperature (Y_1) and slag temperature (Y_2) in metallurgical engineering are inferior to the original RF-LSSVM and RF-RVM algorithms.

Table 8. Influence of matte and slag temperature on the accuracy of RF-LSSVM and RF-RVM.

Indicators	MAE		RMSE		MAPE		R^2	
Model	RF-LSSVM	RF-RVM	RF-LSSVM	RF-RVM	RF-LSSVM	RF-RVM	RF-LSSVM	RF-RVM
Y_1	7.78	7.88	9.92	9.94	14.50%	15.00%	0.93	0.92
Y_2	10.47	10.70	13.37	13.38	17.50%	18.00%	0.94	0.93

When predicting the matte temperature (Y_1), the MAE of the RF-LSSVM model was 7.78, which is lower than that of the RF-RVM model at 7.88. Additionally, the RMSE was lower than that of the RF-RVM model with an MAPE of 15.00%. Furthermore, the R^2 was 0.93, exceeding the

R^2=0.92 of the RF-RVM model. In predicting the slag temperature (Y_2), the MAE of the RF-LSSVM model was 10.47, lower than that of the RF-RVM at 10.70. The RMSE for the RF-LSSVM model was 13.37, slightly lower than that of the RF-RVM model. The MAPE of the RF-LSSVM was lower than that of the RF-RVM model. The R^2 value for the RF-LSSVM model was 0.94, which is higher than that of the RF-RVM at 0.93. Overall, it is evident that the RF-LSSVM model demonstrates superior predictive performance compared with the RF-RVM model when considering the four main control factors. Moreover, when comparing the prediction results of the RF-LSSVM model and the RF-RVM model based on the four main control factors, it is found that for matte temperature (Y_1), the MAE and RMSE of the RF-LSSVM model based on the four main control factors are 1.29% and 0.20% higher than those of the RF-RVM model based on the four main control factors, respectively. For slag temperature (Y_2), the MAE and RMSE of the RF-LSSVM model based on the four main control factors are 2.20% and 0.07% higher than those of the RF-RVM model based on the four main control factors, respectively. It can be seen that the accuracy of the RF-LSSVM algorithm based on the four main control factors is still higher than that of the RF-RVM. Overall, the RF-LSSVM algorithm fully utilizes the advantages of the RF algorithm in screening the main control factors for matte temperature (Y_1) and slag temperature (Y_2). It eliminates redundant information among factors such as the feeding amount (X_1), oxygen pressure (X_2), oxygen flow (X_3), pipeline pressure (X_4), lance windpipe back pressure (X_5), lance windpipe flow (X_6), total air flow (X_7), oxygen concentration (X_8), and exhaust gas residual oxygen concentration (X_9).

4. Conclusions

In this work, through the analysis and control of the main control factors in the molten pool smelting process, the importance of various influencing factors on the smelting quality was determined. The main conclusions are as follows.

(1) RF was used to screen out features that were highly correlated with the predicted slag temperature data, including the oxygen pressure and oxygen flow rate. This process effectively reduces the complexity of the prediction model by eliminating redundant data.

(2) A hybrid model, RF-LSSVM, which integrates RF and the LSSVM, is proposed. This model combines the feature selection capabilities of RF with the nonlinear modeling capabilities of the LSSVM, significantly enhancing the accuracy and robustness of predictions for matte temperature and slag temperature.

(3) This work proposed the RF-LSSVM system. This system absorbs the advantages of a single model. For matte temperature prediction, the MAE reaches 7.87, RMSE reaches 9.95, MAPE reaches 15.70%, and R^2 reaches 0.93. Meanwhile, for slag temperature prediction, MAE reaches 10.19, RMSE reaches 13.11, MAPE reaches 17.45%, and R^2 reaches 0.94.

However, their practical application in industrial settings may face several challenges, which require careful consideration, such as how to ensure the quality of date acquired. Industrial environments often suffer from incomplete or noisy data due to sensor malfunctions or harsh operating conditions. We should implement robust data preprocessing techniques, such as outlier detection and imputation methods, to ensure the data's reliability. Additionally, integrating IoT-based sensors can enhance the efficiency of data collection.

In future work, the model proposed in this work will be applied for the prediction of copper smelting temperature and flue gas oxygen concentration in oxygen-rich top-blowing furnaces to validate the generalization ability of the RF-LSSVM model. Additionally, by combining RF-LSSVM with multi-objective optimization algorithms, a comprehensive framework is developed to achieve

synergistic optimization of temperature control alongside energy consumption, product quality, and other production objectives. At the same time, an interactive visualization interface has been developed to intuitively display the importance of features, prediction results, and error analysis. This interface assists production staff in gaining a deeper understanding of the model and enhances confidence in and the usability of the model.

Author contributions

Senyuan Yang: Methodology, software, writing – original draft preparation; Bo Yu: Methodology; Jianxin Pan: Improved algorithm; Wuliang Yin: Methodology; Hua Wang: Supervision; Kai Yang: Data curation, methodology, writing – original draft preparation; Qingtai Xiao: Writing – reviewing and editing. All authors have read and approved the final version of the manuscript for publication.

Use of Generative AI tools declaration

The authors declare they have not used artificial intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors acknowledge the financial support from the Yunnan Fundamental Research Project, China (No. 202201BE070001-026); the Scientific and Technological Talent and Platform Project of Yunnan Province, China (No. 202405AF140068); the Yunnan Major Scientific and Technological Project (No. 202302AQ370001-4); the Central Guidance Local Scientific and Technological Development Funds (No. 202407AB110022); and the Young Elite Scientist Sponsorship Program of China Association for Science and Technology, China (No. YESS20210106).

Conflict of interest

All authors declare no conflicts of interest in this paper.

Prof. Qingtai Xiao is the Guest Editor of special issue "Advanced in Engineering Statistics, Technology, and Applications" for AIMS Mathematics. Prof. Qingtai Xiao was not involved in the editorial review and the decision to publish this article.

References

1. Y. Wang, Z. Wen, H. Li, Simulation of flows of hazardous elements in copper smelting process based on Bayesian network, *J. Clean. Prod.*, **380** (2022), 135137. https://doi.org/10.1016/j.jclepro.2022.135137
2. D. K. Choshnova, B. S. Stefanov, Study of the conditions of the formation of circulation zones in a laboratory model of a flash smelting furnace, *Russ. J. Non-Ferr. Met.*, **51** (2010), 434–438. https://doi.org/10.3103/S1067821210050081
3. G. Wang, J. Zhang, Y. Wang, Y. Tan, Z. Li, B. Zhang, et al., Study on the bath smelting reduction reaction and mechanism of iron ore: A review, *Metals*, **13** (2023), 672. https://doi.org/10.3390/met13040672

4. X. Y. Chen, J. Y. Dai, Y. S. Luo, Temperature prediction model for a regenerative aluminum smelting furnace by a just-in-time learning-based triple-weighted regularized extreme learning machine, *Processes*, **10** (2022), 1972. https://doi.org/10.3390/PR10101972
5. K. Feng, L. Z. Yang, B. X. Su, W. Feng, L. F. Wang, An integration model for converter molten steel end temperature prediction based on Bayesian formula, *Steel Res. Int.*, **93** (2021). https://doi.org/10.1002/SRIN.202100433
6. O. Tadrari, M. Lacroix, Prediction of protective banks in high temperature smelting furnaces by inverse heat transfer, *Int. J. Heat Mass Tran.*, **49** (2006), 2180–2189. https://doi.org/10.1016/j.ijheatmasstransfer.2005.11.023
7. Q. Shi, J. Tang, M. S. Chu, Process metallurgy and data-driven prediction and feedback of blast furnace heat indicators, *Int. J. Min. Met. Mater.*, **31** (2024), 1228–1240. https://doi.org/10.1007/S12613-023-2693-7
8. W. Xu, J. J. Liu, J. M. Li, H. Wang, Q. T. Xiao, A novel hybrid intelligent model for molten iron temperature forecasting based on machine learning, *AIMS Math.*, **9** (2024), 1227. https://doi.org/10.3934/math.2024061
9. B. Zhao, J. X. Zhao, W. Wu, F. Zhang, T. L. Yao, Research on prediction model of converter temperature and carbon content based on spectral feature extraction, *Sci. Rep.*, **13** (2023), 14409. https://doi.org/10.1038/S41598-023-41751-9
10. X. Liu, Y. P. Bao, L. H. Zhao, C. Gu, Establishment and application of steel composition prediction model based on t-distributed stochastic neighbor embedding (t-SNE) dimensionality reduction algorithm, *J. Sustain. Metall.*, **10** (2024), 509–524. https://doi.org/10.1007/S40831-024-00798-2
11. P. F. Tan, D. Neuschütz, A thermodynamic model of nickel smelting and direct high-grade nickel matte smelting processes: Part I. Model development and validation, *Metall. Mater. Trans. B*, **32** (2001), 341–351. https://doi.org/10.1007/s11663-001-0057-z
12. S. C. Koria, M. K. Barui, L. K. Pandey, Thermochemical model of a 2-stage smelting reduction process for ironmaking, *Scand. J. Metall.*, **28** (1999), 17–24. https://doi.org/10.1007/s11663-001-0057-z
13. K. Yang, Y. L. Wang, M. Wang, J. X. Pan, H. Wang, Q. T. Xiao, A unified heat transfer model for gas-liquid two-phase mixing process in a rectangular channel based on steady status identification, *Appl. Therm. Eng.*, **236** (2024), 121612. https://doi.org/10.1016/J.APPLTHERMALENG.2023.121612
14. K. Yang, X. X. Zhang, G. Y. Yang, M. Li, H. Wang, Q. T. Xiao, Dynamic chaos of imaging measurements for characterizing gas-liquid nonlinear flow behavior in a metallurgical reactor stirred by top-blown air, *Can. J. Chem. Eng.*, **102** (2024), 979–995. https://doi.org/10.1002/CJCE.25077
15. D. Senthilkumar, D. G. Washington, A. K. Reshmy, M. Noornisha, Multi-task learning framework for predicting water quality using non-linear machine learning technique, *J. Intell. Fuzzy Syst.*, **42** (2022), 5667–5679. https://doi.org/10.3233/JIFS-212117
16. D. Mora-Mariano, A. Flores-Tlacuahuac, A machine learning approach for the surrogate modeling of uncertain distributed process engineering models, *Chem. Eng. Res. Des.*, **186** (2022), 433–450. https://doi.org/10.1016/J.CHERD.2022.07.050
17. G. L. Shen, M. F. Li, J. L. Lin, J. Bao, T. He, An empirical study for adopting machine learning approaches for gas pipeline flow prediction, *Math. Probl. Eng.*, 2020, 1–13. https://doi.org/10.1155/2020/7842847

18. G. B. Zou, J. W. Zhou, T. Song, J. W. Yang, K. Li, Hierarchical intelligent control method for mineral particle size based on machine learning, *Minerals*, **13** (2023), 1143. https://doi.org/10.3390/MIN13091143
19. J. Grothoff, N. C. Torres, T. Kleinert, Assessment of reinforcement learning applications for industrial control based on complexity measures, *Au-Autom.*, **70** (2022), 53–66. https://doi.org/10.1515/AUTO-2021-0118
20. H. W. Zhang, Z. Q. Ge, L. Z. Ye, Z. H. Song, Vision-based fan speed control system in the copper scraps smelting process, *Asian J. Control*, **17** (2015), 1742–1755. https://doi.org/10.1002/asjc.996
21. F. Chang, J. Yang, H. L. Lu, H. X. Li, A LIBS quantitative analysis method for samples with changing temperature via functional data analysis, *J. Anal. Atom. Spectrom.*, **36** (2021), 1007–1017. https://doi.org/10.1039/D0JA00514B
22. W. J. Kong, J. L. Ding, Online learning algorithm for LSSVM based modeling with time-varying kernels, *IFAC-Pap.*, **51** (2018), 626–630. https://doi.org/10.1016/j.ifacol.2018.09.354
23. A. Seif, M. Hafezi, C. Jarzynski, Machine learning the thermodynamic arrow of time, *Nature Pub. Group*, **17** (2021), 105–113. https://doi.org/10.1038/S41567-020-1018-2
24. T. H. Zhang, X. Q. Guo, H. Zheng, Y. Liu, A. Wulamu, H. Chen, et al., Review on perovskite-type compound using machine learning, *Sci. Adv. Mater.*, **14** (2022), 1001–1017. https://doi.org/10.1166/SAM.2022.4302
25. K. Yang, B. Yu, W. L. Yin, M. Wang, H. Wang, Q. T. Xiao, Investigation on spatter characteristics of liquid phase and life span of submerged lance in the top-blown smelting process using hydraulic modelling, *Adv. Powder Technol.*, **35** (2024), 104492. https://doi.org/10.1016/J.APT.2024.104492
26. E. Assareh, A. Riaz, M. Ahmadinejad, S. Hoseinzadeh, M. Z. Abdehvand, S. Keykhah, et al., Enhancing solar thermal collector systems for hot water production through machine learning-driven multi-objective optimization with phase change material (PCM), *J. Energy Storage*, **73** (2023), 108990. https://doi.org/10.1016/J.EST.2023.108990
27. K. Hareharen, T. Panneerselvam, R. R. Mohan, Improving the performance of machine learning model predicting phase and crystal structure of high entropy alloys by the synthetic minority oversampling technique, *J. Alloy. Compd.*, **991** (2024), 174494, https://doi.org/10.1016/j.jallcom.2024.174494
28. K. B. Barasu, M. Kothandaraman, Survey of deep learning paradigms for speech processing, *Wireless Pers. Commun.*, **125** (2022), 1913–1949. https://doi.org/10.1007/S11277-022-09640-Y
29. L. J. Shen, Q. Qian, A virtual sample generation algorithm supporting machine learning with a small-sample dataset: A case study for rubber materials, *Comp. Mater. Sci.*, **211** (2022), 111475. https://doi.org/10.1016/J.COMMATSCI.2022.111475
30. C. Liu, L. X. Tang, C. C. Zhao, A novel dynamic operation optimization method based on multiobjective deep reinforcement learning for steelmaking process, *IEEE T. Neur. Net. Lear.*, **35** (2024), 3325–3339. https://doi.org/10.1109/TNNLS.2023.3244945
31. B. Yang, L. Liu, H. Huang, Y. Wang, D. Li, Q. Yang, et al., A real-time temperature field prediction method for steel rolling heating furnaces based on graph neural networks, *Int. J. Heat Mass Tran.*, **235** (2024), 126220. https://doi.org/10.1016/J.IJHEATMASSTRANSFER.2024.126220
32. Z. Y. Ji, W. H. Tao, L. X. Zhang, A boiler oxygen content and furnace temperature prediction model based on honey badger algorithm optimized neural network, *Eng. Res. Express*, **6** (2024). https://doi.org/10.1088/2631-8695/AD22BE

33. Y. Z. Tan, W. Xu, K. Yang, S. Pasha, H. Wang, M. Wang, et al., Predicting cobalt ion concentration in hydrometallurgy zinc process using data decomposition and machine learning, *Sci. Total Environ.*, **962** (2025), 178420. https://doi.org/10.1016/J.SCITOTENV.2025.178420
34. L. S. Chen, Y. M. Wu, Y. B. Liu, T. S. Liu, X. J. Sheng, Time-series prediction of iron and silicon content in aluminium electrolysis based on machine learning, *IEEE Access*, **9** (2021), 10699–10710. https://doi.org/10.1109/ACCESS.2021.3050548
35. K. Lee, J. T. Rinker, C. Tan, M. M. Pour, P. Geng, E. B. Carlson, et al., Unveiling the correlation between weld structure and fracture modes in laser welding of aluminum and copper using data-driven methods, *J. Mater. Process. Tech.*, **338** (2025), 118752. https://doi.org/10.1016/J.JMATPROTEC.2025.118752
36. R. A. John, N. Yantara, S. E. Ng, M. I. Patdillah, M. R. Kulkarni, N. F. Jamaludin, et al., Diffusive and drift halide perovskite memristive barristors as nociceptive and synaptic emulators for neuromorphic computing, *Adv. Mater.*, **33** (2021). https://doi.org/10.1002/adma.202007851
37. Y. H. Li, R. J. Zhu, Y. Q. Wang, L. Y. Feng, Y. Liu, Center-environment deep transfer machine learning across crystal structures: from spinel oxides to perovskite oxides, *NPJ Comput. Mater.*, **9** (2023). https://doi.org/10.1038/S41524-023-01068-7
38. T. Wang, M. Q. Shao, R. Guo, F. Tao, G. Zhang, H. Snoussi, et al., Surrogate model via artificial intelligence method for accelerating screening materials and performance prediction, *Adv. Funct. Mater.*, **31** (2021). https://doi.org/10.1002/adfm.202006245

AIMS Press