



---

*Research article*

## A novel hybrid intelligent model for molten iron temperature forecasting based on machine learning

Wei Xu<sup>1,2</sup>, Jingjing Liu<sup>3</sup>, Jinman Li<sup>4,\*</sup>, Hua Wang<sup>1,2</sup> and Qingtai Xiao<sup>1,2,\*</sup>

<sup>1</sup> State Key Laboratory of Complex Nonferrous Metal Resources Clean Utilization, Kunming University of Science and Technology, Kunming 650093, China

<sup>2</sup> Faculty of Metallurgical and Energy Engineering, Kunming University of Science and Technology, Kunming 650093, China

<sup>3</sup> School of Mathematics, Science and Engineering, University of the Incarnate Word, San Antonio, Texas 78209, United States

<sup>4</sup> Department of Quality Management, Inspection and Testing, Yibin University, Yibin, 644000, China

\* **Correspondence:** Email: [kmustmz@126.com](mailto:kmustmz@126.com), [qingtai.xiao@kust.edu.cn](mailto:qingtai.xiao@kust.edu.cn).

**Abstract:** To address the challenges of low accuracy and poor robustness of traditional single prediction models for blast furnace molten iron temperature, a hybrid model that integrates the improved complete ensemble empirical mode decomposition with adaptive noise, kernel principal component analysis, support vector regression and radial basis functional neural network is proposed for precise and stable iron temperature prediction. First, the complete ensemble empirical mode decomposition is employed to decompose the time series of iron temperature, yielding several intrinsic mode functions. Second, kernel principal component analysis is used to reduce the dimensionality of the multi-dimensional key variables from the steel production process, extracting the major features of these variables. Then, in conjunction with the K-means algorithm, support vector regression is utilized to predict the first column of the decomposed sequence, which contains the most informative content, evaluated using the Pearson correlation coefficient method and permutation entropy calculation. Finally, radial basis function neural network is applied to predict the remaining time series of iron temperature, resulting in the cumulative prediction. Results demonstrate that compared to traditional single models, the mean absolute percentage error is reduced by 54.55%, and the root mean square error is improved by 49.40%. This novel model provides a better understanding of the dynamic temperature variations in iron, and achieves a hit rate of 94.12% within a range of  $\pm 5^{\circ}\text{C}$ . Consequently, this work offers theoretical support for real-time control of blast furnace molten iron temperature and

holds practical significance for ensuring the stability of blast furnace smelting and implementing intelligent metallurgical processes.

**Keywords:** molten iron temperature; intelligent prediction; K-means; empirical mode decomposition; machine learning

**Mathematics Subject Classification:** 62P30

---

## 1. Introduction

With the rapid industrialization of the world, energy consumption in industry has been increasing and inefficient use of energy consumption and waste of limited resources has led to excessive carbon emissions [1]. The ferrous metals manufacturing industry (i.e., steel industry) is considered one of the highest energy-consuming and carbon-emitting sectors, and it is also one of the largest energy-consuming manufacturing industries globally [2]. This industry heavily relies on coal and other fossil fuels for production, resulting in substantial emissions of carbon dioxide, accounting for approximately 4% to 5% of the total global anthropogenic carbon emissions. According to the reports from the International Energy Agency, the carbon dioxide (i.e., CO<sub>2</sub>) emissions from the manufacturing industry account for 40% of world carbon dioxide emissions, with the steel manufacturing sector being the largest contributor, making up around 27% of the global manufacturing sector [3].

In terms of the metallurgical industry, the process of iron smelting in blast furnaces is one of the crucial steps in the reduction and production process. The physical temperature of molten iron constitutes approximately 50% of the total heat input during the entire process. Therefore, effectively controlling the blast furnace molten iron temperature within a reasonable range is essential for reducing the energy consumption in the reduction smelting process [4]. In other words, temperature control is one of the important indicators of product quality [5,6]. To achieve stable pig iron quality and consistent blast furnace performance, the temperature of hot metal should be maintained within a stable range. The stable temperature inside the furnace determines the composition and quality of hot metal, ensuring the smooth operation of the blast furnace [7]. Besides ensuring the highest product quality, reducing production costs is vital for the economic viability of enterprises. Hence, the establishment of a closed-loop process model that accurately predicts the status of the system could significantly reduce production costs [8]. However, directly measuring the blast furnace molten iron temperature is often a challenging task due to the complex dynamics and high-temperature characteristics of the smelting process [9,10].

Due to the complex combustion processes and interactions among multiple key variables inside the blast furnace, the iron temperature exhibits nonlinear and non-steady-state characteristics [11]. Therefore, establishing a mechanistic model for accurately predicting the iron temperature is challenging, especially due to the lack of high-temperature thermodynamic and kinetic data, as well as the intricate dependencies among internal system variables and the three target variables under study [12]. In the past, expert systems were considered promising alternative solutions for blast furnace modeling. Some groundbreaking work has been accomplished in the field of expert systems [13,14]. Moreover, expert rules have been embedded in artificial and logical intelligent systems to supervise the process adequately. Expert rules provide temporary assumptions for each task under supervision, such as gas flow distribution or furnace temperature levels, followed by appropriate actions based on the outcomes of these temporary assumptions. However, these methods often overlook the complex

interactions among multiple key variables, leading to unstable model predictions [15]. In recent years, data-driven methods have gradually gained prominence in blast furnace modeling and process state evaluation. For instance, Jimenez and colleagues have employed a back propagation neural network (BP-NN) model to predict blast furnace iron temperature [16].

As the furnace temperature rises, the silicon content in molten iron gradually increases, and there is an approximately linear correlation between the silicon content in molten iron and temperature. Therefore, attributes closely related to iron temperature have selected as influencing factors from some studies [17,18] and data-driven algorithms have been employed to investigate the silicon content in molten iron and predict the molten iron temperature [19]. However, due to over-reliance on a single parameter and the high lag in silicon content in molten iron, there is necessity for further improvement in model stability. For instance, Su et al. [20] and Zhang et al. [21] have employed an extreme learning machine (ELM) for molten iron temperature prediction. In comparison to other neural networks, radial basis function neural networks (RBF-NN) exhibit the capability to approximate any nonlinear function with arbitrary precision, showcasing superior approximation, classification and global optimality. Furthermore, RBF-NN is local approximation network that omits the learning behavior of hidden layer weights, thus avoiding the time-consuming process of layer-by-layer error propagation. As a result, RBF-NN exhibits a rapid learning convergence speed. RBF-NN finds extensive applications in various simulation processes within the chemical industry [22,23]. Due to its effectiveness in handling non-linear relationships and its strong capabilities in dealing with high-dimensional data, the sample size requirement is not high [24]. Support vector machine (SVR) is suitable for predicting multi-dimensional related variables. In fact, SVR has been applied in complex, non-linear data modeling and prediction tasks including blast furnace modeling [25,26].

To address the issue of irregularity in the iron temperature time series, decomposition theory has been employed to process the original sequences, achieving promising results in predicting smelting temperatures [27], wind speeds [28] and fluid properties [29], among other areas. For instance, Wang et al. decomposed the hybrid mode of data, then predicted the decomposed components and finally reconstructed the predicted components to predict high-noise, non-stationary and nonlinear short-term CO<sub>2</sub> emissions with the same nature as hot metal temperature, and provided quantitative benchmarks for policy formulation [30]. By decomposing the original sequences, useful information could be successfully extracted. Notably, the improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) method decomposes sequences into intrinsic mode functions (IMFs) with different frequencies and energy features, offering a better grasp of the details and changing trends in temperature signals with different frequency components. By extracting relevant features from temperature data, the method reduces the amount of decomposition, enhancing prediction accuracy and training speed and making the decomposition more accurate and interpretable. Therefore, it has gained widespread application in decomposition algorithms.

Furthermore, to address the issue of different modes and multi-layered trends in iron temperature, Cui et al. used the fuzzy C-means model to divide the parameter set of the integrated furnace temperature characterization, predicted the furnace parameters step by step to achieve dynamic prediction of hot metal silicon content, and adjusted the control parameters reversely to achieve accurate control of furnace temperature [31]. Fontes et al. applied fuzzy C-means to cluster similar data objects and separate different data objects, which improved the performance of the model [32]. Zhao et al. adopted the simple and easy K-means algorithm, and finally, the performance of the prediction model of molten iron temperature based on the K-means clustering label was further improved, and the accurate monitoring and prediction of molten iron temperature was realized [33]. Clustering analysis including K-means and other algorithms, known for their rapid convergence,

minimal hyperparameter adjustments and strong interpretability, has been used to group data objects based on certain attributes, forming multiple classes or clusters, i.e., different feature groups. This ensures the maximization of data object similarity within the same class or cluster and the maximization of differences between different classes or clusters. Subsequently, data within each cluster would be individually trained for corresponding predictions. In summary, decomposition theory and clustering concepts provide effective methods to address the challenges of iron temperature time series and the diverse modes and trends of iron temperature. A range of data analysis tools and techniques have been reported to meet diverse and complex needs.

While decomposition algorithms and clustering techniques contribute to a more thorough comprehension of features and trends in molten iron temperature data, potential issues arise from the IMFs generated by ICEEMDAN and the results of K-means clustering, particularly concerning the number of components and clusters. To optimize model efficiency and generalization capabilities, the kernel principal component analysis (KPCA) algorithm could be opted to efficiently reduce dimensions when handling high-dimensional data, mitigating the risk of model complexity induced by an excess of components and clusters [34]. In fact, the reduction in data dimensions facilitates the precise capture of critical features in molten iron temperature data, concurrently enhancing the model generalization prowess. Furthermore, KPCA excels in handling non-linear relationships, making it well-suited for addressing potential complex non-linear relationships within molten iron temperature data. This is pivotal for ensuring accurate and robust predictions. The decision to employ the KPCA algorithm is strategic, aiming to strike a balance in handling molten iron temperature data, optimizing model efficiency, streamlining structure and ensuring a comprehensive understanding of data features and trends.

In this work, this molten iron temperature prediction model utilizes various sets of data collected from the steel production process. It applies decomposition theory (i.e., ICEEMDAN) to decompose the molten iron temperature data, employs KPCA method for dimensionality reduction and utilizes clustering regression to predict IMF, which contains the highest information content. For the remaining sequences with a stronger temporal aspect, the RBF-NN is employed for training and prediction. The final step involves aggregating and reconstructing these results to obtain the ultimate prediction value. The contribution of this work is that the ICEEMDAN is introduced to decompose the molten iron temperature series to deal with the problem of nonlinearity and disorder of the molten iron temperature series for the first time. Furthermore, the laws of IMF components that pass through ICEEMDAN are explored, and K-means, SVR and RBF-NN are used to predict the components with different laws, respectively. Overall, this model provides valuable guidance for on-site operators, aiding in stabilizing furnace operations and achieving real-time furnace control. The work not only guides the field of metallurgical and chemical industries but also may be applied to other nonlinear sequence prediction fields.

The rest of this article is organized as follows. In Section 2, the experiment and related methods are provided. In Section 3, The decomposition algorithm and the processing results of related parameters are introduced, as well as the results and related indicators of the proposed model. In Section 4, the conclusions are provided, along with an outlook for future work.

## 2. Methods

### 2.1. Related algorithms

#### 2.1.1. Data decomposition

Traditional methods for forecasting molten iron temperature often struggle to accurately predict abrupt changes, resulting in less precise model predictions. Decomposing molten iron temperature into multiple components with various frequencies and building separate forecasting models for each component is a highly effective and feasible solution. Among these methods, complete ensemble empirical mode decomposition with adaptive noise is an improvement over empirical mode decomposition. It effectively overcomes issues like modes mixing that occur after empirical mode decomposition. However, even with complete ensemble empirical mode decomposition with adaptive noise, the components obtained may still contain noise and pseudo modes. To address this, an improved approach (i.e., ICEEMDAN) is introduced. In fact, the  $k$ -th mode component IMF is derived from empirical mode decomposition with added white noise but not direct addition of Gaussian white noise during the traditional decomposition process. The specific steps of the ICEEMDAN are as follows.

(1) Incorporate a set of white noise  $w^i(t)$  into the molten iron temperature sequence  $x(t)$ , where  $x_1^i(t)$  is constructed as

$$x_1^i(t) = x(t) + \zeta_0 E_1[w^i(t)]$$

to generate the sequence corresponding to this set a white noise  $r_1(t)$ . Subsequently, the first residual component  $r_1(t)$  is obtained by

$$r_1(t) = \langle N(x_1^i(t)) \rangle, \quad (1)$$

where  $N(x_1^i(t))$  represents the calculation of the local mean of the data sequence  $x_1^i(t)$ .

(2) The IMF is computed by

$$\text{IMF}_1 = x(t) - r_1(t) = x(t) - \langle N(x_1^i(t)) \rangle, \quad (2)$$

where  $\text{IMF}_1$  refer as to the first IMF.

(3) Construct the sequence

$$x_2^i(t) = x(t) + \zeta_0 E_2[w^i(t)]$$

resulting in the second residual component

$$r_2(t) = \langle N(x_2^i(t)) \rangle.$$

The second intrinsic mode function  $\text{IMF}_2$  is calculated by

$$\text{IMF}_2 = x(t) - r_2(t) = x(t) - \langle N(x(t) + \zeta_0 E_2[w^i(t)]) \rangle, \quad (3)$$

where  $E_2(\cdot)$  represents the second IMF which is generated by the empirical mode decomposition of white noise  $w^i(t)$ ,  $\zeta_1$  is the signal-to-noise ratio for the added noise divided by the standard deviation of this noise component  $E_2(\cdot)$ , and  $N(x_2^i(t))$  denotes the calculation of the local mean of the data sequence  $x_2^i(t)$ .

(4) Repeating the above step, the corresponding residual component and the IMF can be obtained

by

$$\begin{cases} r_k(t) = \langle N(r_k(t) + \zeta_{k-1} E_k[w^i(t)]) \rangle, \\ \text{IMF}_k = r_{k-1}(t) - r_k(t), \end{cases} \quad (4)$$

where  $r_k(t)$  refers as the  $k$ -th residual component and  $\text{IMF}_k$  refers as the  $k$ -th IMF.

### 2.1.2. Dimension reduction

KPCA would be a technique used for dimensionality reduction of temperature data and capable of handling nonlinear relationships within temperature data. It has been considered as a nonlinear extension of the principal component analysis method. Its core idea involves mapping the original temperature data through a nonlinear function into a high-dimensional space, where the primary features of the temperature data are identified. In terms of the applications for predicting molten iron temperature, KPCA could help operators better understand complex temperature data and improves the performance of predictive models. The computation procedure of KPCA is provided as follows.

(1) The covariance matrix of temperature data mapped in the high-dimensional space is given by

$$\text{cov} = \frac{1}{N} \sum_{i=1}^N \phi(x_i)(x_i)^T, \quad (5)$$

where  $\text{cov}$  refers as the covariance matrix.

(2) By applying the above equation, the characteristic equation could be derived by

$$\lambda \cdot w = \text{cov} \cdot w, \quad (6)$$

where  $\lambda$  represents the eigenvalue, and  $w$  is the corresponding eigenvector.

(3) With the help of transforming, the above equation can be rewritten by

$$\lambda \cdot [\phi(x_j), w] = [\phi(x_j), \text{cov} \cdot w], \quad (7)$$

where  $w$  can be replaced with

$$w = \sum_{i=1}^N v_i \phi(x_i). \quad (8)$$

(4) Thus, the above equation can be rewritten by

$$\lambda \sum_{i=1}^N v_i [\phi(x_j), \phi(x_i)] = \frac{1}{N} \sum_{i=1}^N v_i [\phi(x_j), \sum_{k=1}^N \phi(x_k)] [\phi(x_k), \phi(x_i)]. \quad (9)$$

(5) By introducing the kernel function, the projection of the mapped temperature data in the high-dimensional space can be calculated as

$$\text{Proj}_i = \sum_{i=1}^N v_i^j \cdot \text{Ker}(x_i, x_j) = \sum_{i=1}^N v_i^j \cdot [\phi(x_i), \phi(x_j)], \quad (10)$$

where the kernel function employed in this paper is Gaussian kernel function.

### 2.1.3. Series prediction

(1) K-means clustering is an iterative data analysis algorithm. Its principle is to divide the data into  $K$  groups and initially select  $K$  objects randomly as cluster centers. The distance between each object and all the cluster centroids was calculated, and each object was assigned to the nearest cluster centroid. The centroids of the clusters and the objects assigned to them represent a cluster. In this work,

there are 64 temperature data points are used for clustering, and  $K$  is defined as 2 according to the amount of temperature data. Use the Euclidean distance to figure out which cluster to put in. The steps to determine the cluster to which a new temperature data point belongs are outlined below. Compute the Euclidean distance between the new temperature data point and the cluster centers. For a new temperature data point, the Euclidean distance between it and each cluster center can be computed. Therefore, the Euclidean distance is computed by

$$d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, \quad (11)$$

where  $d$  is the Euclidean distance,  $n$  represents the temperature data point dimensionality, indicating the number of features in the temperature data point,  $x_i$  stands for the  $i$ -th feature value of the new temperature data point and  $y_i$  denotes the  $i$ -th feature value of the cluster center. Subsequently, the new temperature data point is assigned to the cluster with the shortest Euclidean distance. Specifically, the Euclidean distance from the new temperature data point to each cluster center could be compared. If the distance to Cluster 1 is shorter than that to Cluster 2, the new temperature data point is assigned to Cluster 1. Conversely, if the distance to Cluster 2 is shorter, it is allocated to Cluster 2.

(2) Support vector machine is a supervised learning method, which can effectively reveal the correlation between the influencing factors and the target value. Here, K-means cluster analysis is combined with SVR to divide the temperature data into clusters with similar characteristics, which facilitates the identification of potential patterns and clusters within the data. Theoretically, SVR is a regression algorithm built upon the principles of support vector machine. Its functional construction is equivalent to solving a quadratic programming problem. The function involved in solving this problem is given by

$$\min_{\alpha, \alpha^*} \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \times (\varphi(x_i)\varphi(x_j)) + \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) - \sum_{i=1}^l y_i(\alpha_i - \alpha_i^*), \quad (12)$$

where

$$0 \leq \alpha_i, \alpha_i^* \leq C$$

and

$$\sum_{i=1}^l (\alpha_i^* - \alpha_i) = 0.$$

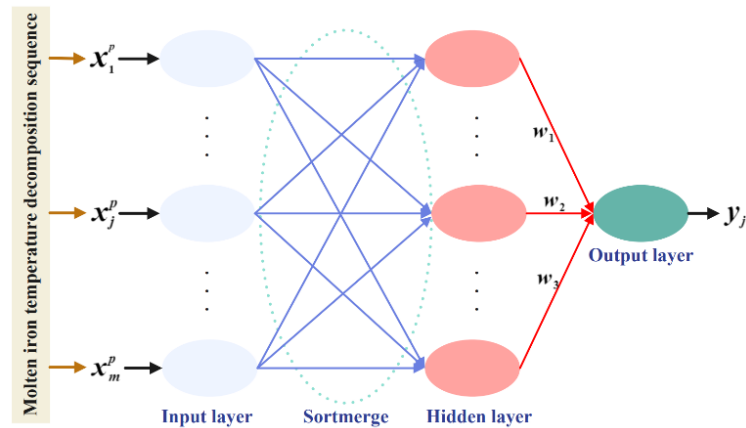
The obtained regression function is given by

$$f(x) = \omega\varphi(x) + b = \sum_{i=1}^l (\alpha_i - \alpha_i^*)K(x_i, x) + b, \quad (13)$$

where  $K(x_i, x)$  is the kernel function. Here, the Gaussian kernel function is given by

$$K(x_i, x_j) = \varphi(x_i)\varphi(x_j) = \exp\left(-\frac{1}{2\sigma^2}|x_i - x_j|^2\right). \quad (14)$$

(3) RBF-NN method which has been primarily used for nonlinear functions is a feedforward network consisting of three layers. The fundamental structure of the RBF-NN comprises three components including the input layer, the radial basis layer (hidden layer) and the output layer, with fully connected nodes between these layers, as depicted in Figure 1.



**Figure 1.** Diagram of basic architecture of RBF-NN used in this work.

Since the hidden layer consists of RBF-NN, it can directly map the input vector into the hidden space, eliminating the need for weighted connections. Consequently, the connection weights between the input layer and the hidden layer are all set to 1. The hidden layer accomplishes the nonlinear projection of the input vector, while the output layer is responsible for the final linear weighted summation. The activation function of RBF-NN can be defined by

$$R(x_p - c_i) = \exp\left(-\frac{1}{2\sigma^2}\|x_p - c_i\|^2\right), \quad (15)$$

where  $\|x_p - c_i\|$  refers as the Euclidean norm,  $c$  refers as the center of the Gaussian function and  $\sigma$  refers as the variance of the Gaussian function. For  $j = 1, 2, \dots, n$ , the output of RBF-NN can be calculated by

$$y_j = \sum_{i=1}^h \omega_{ij} \exp\left(-\frac{1}{2\sigma^2}\|x_p - c_i\|^2\right), \quad (16)$$

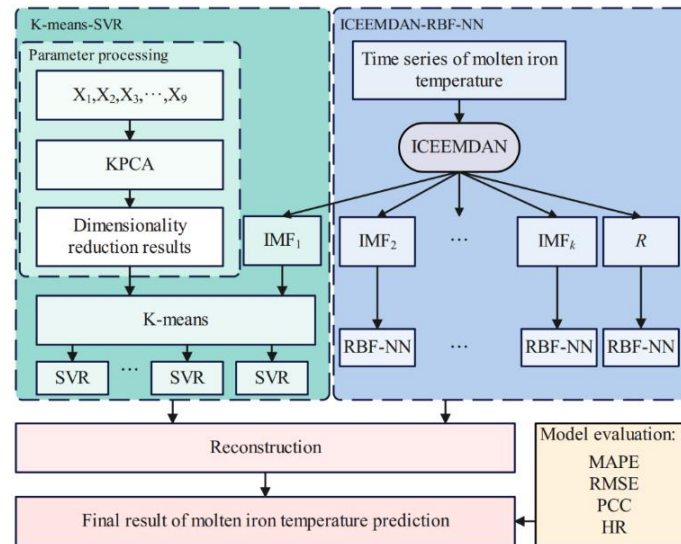
where  $x_p = (x_1^p, x_2^p, \dots, x_m^p)^T$  refers as the  $p$ -th input sample,  $c_i$  refers as the center of the node in the hidden layer of the network,  $\omega_{ij}$  refers as the weight from the hidden layer to the output layer,  $i = 1, 2, \dots, h$  refers as the number of nodes in the hidden layer and  $y_j$  refers as the  $j^{\text{th}}$  actual output. Let  $d$  be the output expectation value of the sample, and the variance  $\sigma$  of the function can be calculated by

$$\sigma = \frac{1}{p} \sum_j^m \left(\|d_j - y_i c_i\|^2\right). \quad (17)$$

## 2.2. Proposed hybrid model

Figure 2 shows the process of establishing the hybrid model of hot blast furnace molten iron temperature. Initially, the molten iron temperature data is decomposed into different frequency components and energy terms, including intrinsic mode functions and a residual component, using the ICEEMDAN method. Subsequently, the relevant parameters derived from KPCA and IMF<sub>1</sub> are employed for clustering. SVR is trained for prediction, and models are applied to predict the remaining IMFs and residual component. The final step involves aggregating and reconstructing the predicted values.





**Figure 2.** Flowchart of the hybrid model proposed in this work.

This new model encompasses the following key steps.

(1) The molten iron temperature and its related data are preprocessed. The parameter data of independent variables were standardized and transformed into a standard normal distribution with mean between 0 and 1 to eliminate the dimensional differences between different parameters. KPCA is used to reduce the dimension of the relevant parameters.

(2) The molten iron temperature sequence is decomposed by ICEEMDAN. ICEEMDAN is employed to decompose the time series of molten iron temperature. The temperature data is extracted as IMFs, and  $R$  is obtained through iteration. This process effectively breaks down the non-stationary and nonlinear complex sequence into a series of more regular sub-sequences.

(3) The molten iron temperature and related parameters are clustered. IMFs and KPCA-reduced relevant parameters are extracted, and a suitable value for  $k$  is selected based on the data characteristics for clustering. Data points are divided into  $k$  clusters.

(4) Model modeling of the processed time series and related data. For each of the  $K$  clusters, SVR models are trained and applied for predicting  $IMF_1$ . As for the remaining IMFs and  $R$  sequences, RBF-NN is employed for temperature prediction.

(5) The performance of the proposed model was evaluated by different evaluation indicators. The results of predictions for individual clusters and sequences are superimposed to reconstruct the overall molten iron temperature prediction. Evaluation metrics including mean absolute percentage error (MAPE), root mean squared error (RMSE), hit ratio (HR) and Pearson correlation coefficient (PCC) are used for assessment and provided as follows.

### 2.3. Evaluation metrics

#### 2.3.1. Model evaluation

To evaluate the model prediction accuracy, the following four error metrics: MAPE, RMSE, HR and PCC are used here and given by

$$\begin{cases} MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \\ RMSE = \frac{1}{N} \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2}, \\ HR = \frac{1}{N} \sum_{i=1}^N \text{hit}(i), \\ PCC = \frac{\text{cov}(y_i, \hat{y}_i)}{\sigma_{y_i} \sigma_{\hat{y}_i}}, \end{cases} \quad (18)$$

where  $y_i$  is the actual molten iron temperature,  $\hat{y}_i$  is the predicted molten iron temperature,  $N$  is the length of the predicted series,  $\text{hit}(i)$  represents the condition in which the predicted value falls within a  $\pm 5^\circ\text{C}$  range of the actual value,  $\text{cov}$  is the covariance operator,  $\sigma_{y_i}$  is the standard deviation of the actual molten iron temperature time series and  $\sigma_{\hat{y}_i}$  is the standard deviation of the predicted molten iron temperature time series. MAPE and RMSE provide quantitative assessments of the absolute extent of errors. Smaller MAPE and RMSE values indicate smaller prediction errors, signifying higher accuracy. MAPE is used to calculate the average percentage error, which represents the average percentage difference between the predicted value and the actual value, and can be compared across different datasets and problems.

Moreover, RMSE considers the square of errors, providing a measurement of the overall error magnitude. Furthermore, HR finds extensive utility in industrial applications, particularly in the realm of molten iron temperature prediction. HR measures a model ability to predict values within a  $\pm 5^\circ\text{C}$  range of actual values. In numerous industrial scenarios, ensuring accuracy within a specific temperature range is of paramount importance, directly impacting the quality and safety of production processes. Last, PCC is employed to gauge the linear relationship between predicted values and actual values. It furnishes information about the strength of the association between two variables, with values closer to 1 or  $-1$  indicating a stronger linear relationship. The comprehensive application of these metrics aids operators or engineers in evaluating various models for molten iron temperature prediction, guiding model selection and refinement, ultimately optimizing industrial processes.

### 2.3.2. Permutation entropy

Permutation entropy would be a nonlinear method for detecting abrupt changes in temperature data and offers a concise way to describe complex systems that were previously challenging to quantify. The magnitude of the permutation entropy value  $H_p(m)$  can represent the complexity of a time series. Time series that are more orderly correspond to smaller values of permutation entropy, while more complex time series correspond to larger permutation entropy values. The calculation process is simple and it exhibits high resistance to interference and robustness. Assuming that  $\{x_i\}$  is a one-dimensional time series, embedded in a dimension is  $m$  with a time delay of  $\tau$ . The sliding window  $(m, \tau)$  that sequentially passes through the current sequence, and the one-dimensional time series would be reconstructed by

$$\mathbf{Y} = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_j & x_{j+\tau} & \cdots & x_{j+(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_k & x_{k+\tau} & \cdots & x_{k+(m-1)\tau} \end{bmatrix}, \quad (19)$$

where  $i = 1, 2, \dots, N$ ,  $m > 1$ ,  $\tau > 0$  and each row of the matrix serves as a reconstruction component is given by

$$Y_j = [x_j, x_{j+\tau}, \dots, x_{j+(m-1)\tau}], \quad (20)$$

where for  $j = 1, 2, \dots, k$ . The reconstruction components are arranged in ascending order based on their numerical values, resulting in a total of  $m!$  possible permutations. Using  $P_1, P_2, \dots, P_k$  to represent the probabilities of occurrence for each permutation, permutation entropy is defined by

$$H_D(m) = -\sum_{i=1}^k P_i \ln(P_i), \quad (21)$$

where the values of  $m$  is selected according to open literature [35].

### 3. Results and discussion

#### 3.1. Data description and pre-processing

In this work, an on-site dataset comprising 80 data points from a metallurgical factory were utilized to validate the effectiveness and accuracy of the proposed hybrid model. The smelting process involved mine, which has a significant impact on the blast furnace operation. As a result, the furnace conditions often exhibit substantial fluctuations, increasing the complexity of the modeling task. Modeling with these data enables effective testing of the model generalization performance. Parameters with a substantial impact on blast furnace temperature were chosen for the investigation. Considering the complexity of the blast furnace smelting environment and the potential for equipment malfunctions affecting data integrity, missing values in the sample data were addressed using the mean imputation method. Due to the relatively limited sample size and the collection of data during stable blast furnace operation, the dataset underwent preprocessing. The numerical ranges of molten iron temperature and related parameters are presented in Table 1.

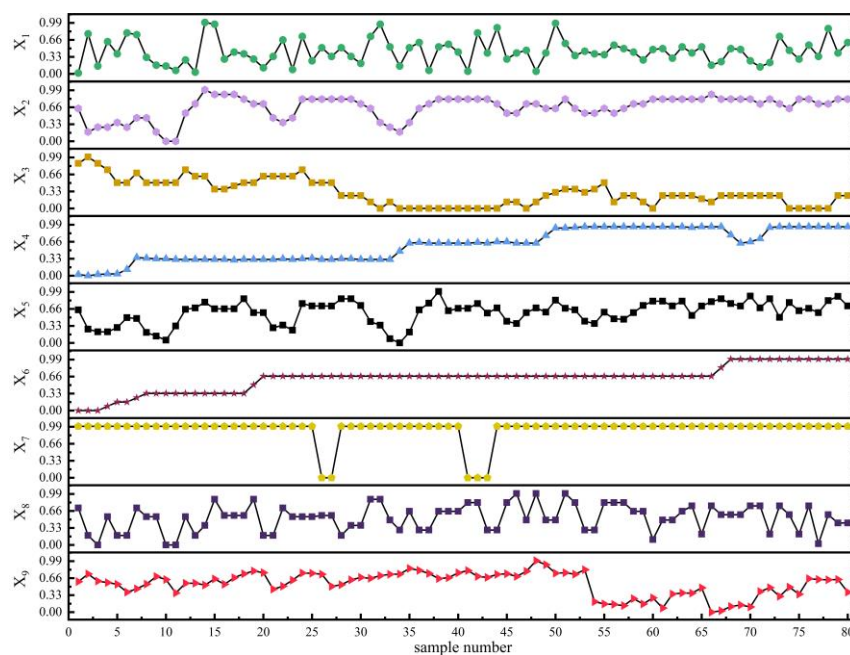
**Table 1.** Character parameters for molten iron temperature prediction.

Variable name	Variable description	Value range	Unit
Y	molten iron temperature	1492.5~1519	°C
X <sub>1</sub>	silicon content	0.43~0.73	%
X <sub>2</sub>	blowing rate	4775~5050	m <sup>3</sup> ·min <sup>-1</sup>
X <sub>3</sub>	coal injection	13~17	kg·t <sup>-1</sup>
X <sub>4</sub>	oxygen enrichment rate	0.79~1.19	%
X <sub>5</sub>	permeability	26585~26294	ml
X <sub>6</sub>	wind temperature	1040~1060	°C
X <sub>7</sub>	blast furnace top pressure	0.184~0.186	kPa
X <sub>8</sub>	comprehensive load	3.10~3.22	kPa
X <sub>9</sub>	gas utilization rate	40.97~46.04	%

After preprocessing, the numerical ranges of the molten iron temperature are 1492.5°C to 1519°C, silicon content ranges from 0.43% to 0.73%, the blowing rate varies between 4775 and 5050 m<sup>3</sup>·min<sup>-1</sup>, coal injection ranges from 13 to 17 kg·t<sup>-1</sup>, oxygen enrichment rate varies between 0.79% and 1.19%, permeability ranges from 26294 to 26585 ml, wind temperature varies between 1040°C and 1060°C, blast furnace top pressure is in the range of 0.184 kPa to 0.186 kPa, comprehensive load is between 3.10 kPa and 3.22 kPa and gas utilization rate varies from 40.97% to 46.04%. These parameters provide

critical data concerning molten iron production and blast furnace operations, enabling the monitoring and control of the smelting process to ensure production stability and quality.

Considering the diverse units and dimensions of the smelting data relevant parameters, the execution of data standardization is a pivotal procedure to ensure the effectiveness of modeling and investigation. The fundamental objective of data standardization is to mitigate the potential impact of varying magnitude among different parameters on the modeling outcomes. The minimum-maximum normalization method is employed to map all parameters onto a uniform scale, typically within the range of 0 to 1. This practice contributes to maintaining consistency in both modeling and investigation. The scatter plot of standardized relevant parameters is depicted in Figure 3. As illustrated in this figure, the scatter plot of standardized sample data for the relevant parameters exhibits a consistent data range. This further enhances our comprehension of the relationships among these parameters, ultimately improving the accuracy and interpretability of our modeling efforts. This standardization process plays a pivotal role in ensuring the scientific rigor and reliability.



**Figure 3.** Normalized line chart of related factors in terms of original data.

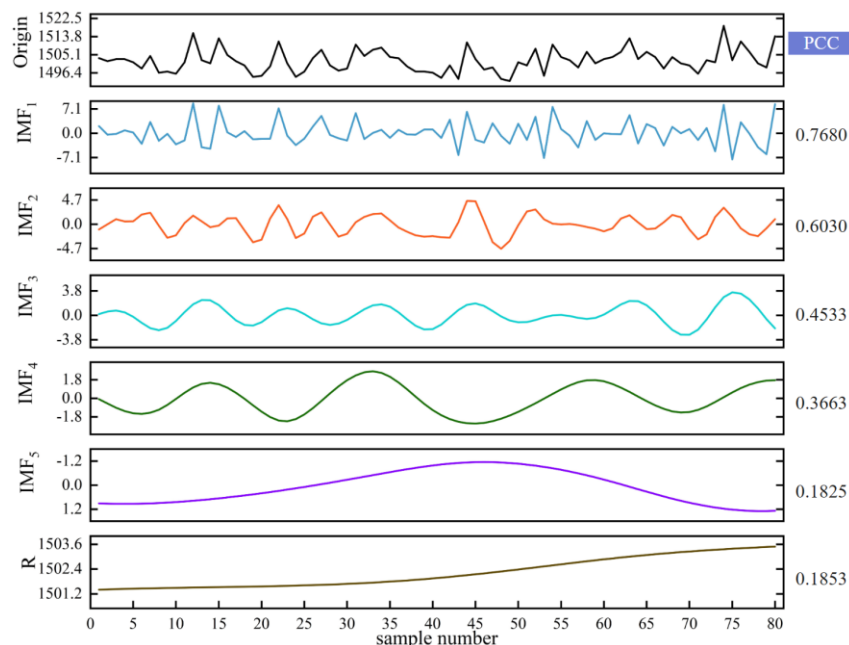
In parallel, to reduce the data dimensionality and enhance computational efficiency, the KPCA method is applied to investigate nine variables (i.e., factors affecting molten iron temperature) within the input blast furnace molten iron model. The results of this investigation are presented in Table 2. Based on two scientific evaluation criteria, variance contribution rate and cumulative variance contribution rate, parameter selection was performed to eliminate less influential feature parameters in predicting. The first seven principal components were identified, with their cumulative variance contribution rate exceeding 90%. This observation signifies that these seven principal components effectively capture most of the dataset information, simultaneously reducing data dimensionality while preserving data richness. As a result, these seven principal components were retained for subsequent investigation and modeling. This selection process contributes to data simplification, reduction of redundancy, improved modeling efficiency and ultimately, ensures that the predefined objectives would be achieved with greater accuracy and efficiency.

**Table 2.** The calculation results of relevant parameters using KPCA method.

No.	Eigenvalues	Eigenvalue variance contribution (%)	Cumulative variance contribution (%)
1	1.26	16.64	16.64
2	1.09	14.33	30.97
3	1.06	13.88	44.85
4	1.00	13.11	57.96
5	0.98	12.89	70.86
6	0.87	11.41	82.27
7	0.70	9.24	91.51
8	0.65	8.49	100.00

### 3.2. Characteristics based on ICEEMDAN

To fully extract the regular features in complex signals to enhance prediction performance, the ICEEMDAN algorithm is utilized to decompose 80 sets of data. During this process, the noise standard deviation is set to 1, the number of realizations is a specific value and the maximum allowed sifting iteration is 5000. Figure 4 presents the decomposition results of the original molten iron temperature sequence, comprising five modal components and one residual component, each characterized by different frequency scales.

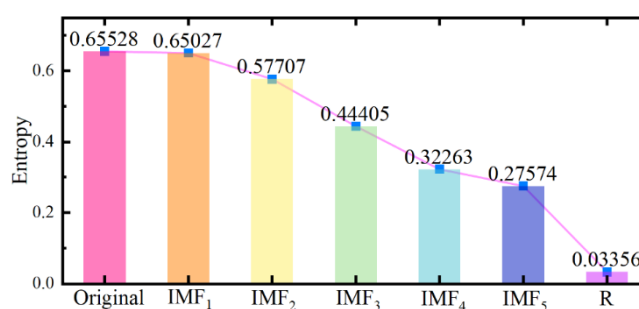
**Figure 4.** Results of iron temperature decomposition using ICEEMDAN.

Additionally, the Pearson correlation coefficients between each modal component and the residual component with the original time series are computed. By the presentation in this figure, the ICEEMDAN method is applied to decompose the temperature dataset, resulting in five IMFs and one residue. This decomposition process exhibits a distinct trend, with the frequency of oscillations

gradually decreasing as more IMFs are added, leading to a gradual smoothing of the data. Findings highlight the excellent performance of the ICEEMDAN decomposition method in capturing local features and periodic variations within the sequence. This implies a better understanding of crucial characteristics within the data, including local peaks and troughs, as well as other nonlinear changes. This decomposition process contributes to providing more accurate and reliable input data for subsequent predictive models. Additionally, it offers a richer data representation that can be utilized for more in-depth investigation and modeling.

IMF<sub>1</sub> demonstrates a notably strong correlation with the original sequence compared to other modal components, with a correlation coefficient as high as 0.7680. This observation reveals that IMF<sub>1</sub> encapsulates highly relevant feature information from the original dataset, encompassing the primary trends, periodic components and key information related to the original molten iron temperature sequence. This emphasizes the pivotal role of IMF<sub>1</sub> in the decomposition process. Due to its close association with the original sequence, IMF<sub>1</sub> proves highly effective in describing the overall dynamic characteristics of the molten iron temperature, thereby enhancing the reliability of modeling for capturing future trends and changes. Furthermore, IMF<sub>1</sub> provides insights into significant signal characteristics within the original temperature data, including primary trends and prominent periodic components. Particularly, for forecasting tasks necessitating an in-depth comprehension of the dynamic properties of molten iron temperature, the relevance of IMF<sub>1</sub> positions it as a critical modeling element, facilitating a better understanding and prediction of the complex behavior of molten iron temperature.

To further investigate the relationship between the decomposed sequences and time series, permutation entropy is employed to quantify the complexity of time series and the molten iron temperature sequence, as illustrated in Figure 5. In the calculation of permutation entropy,  $m=5$  and  $t=1$ . This investigation serves to explore the latent associations between these sequences, providing us with additional insights for a more comprehensive understanding of the dynamic characteristics of molten iron temperature. By observing this figure, it can be noticed that as the number of modal components increases, the permutation entropy values continuously decrease, reaching 0.27574 for the final modal component IMF<sub>5</sub> and 0.03356 for the residual. In contrast, the permutation entropy of IMF<sub>1</sub> is 0.65027, and the permutation entropy of the original molten iron temperature sequence is 0.65528, with little significant difference between them.



**Figure 5.** Arrangement entropy values for the original molten iron temperature sequence and the decomposed sequences.

This indicates that, during the decomposition process, with the increasing IMFs, the complexity of the sequence gradually decreases, as reflected in the decreasing trend of permutation entropy values. This trend suggests that in the initial IMFs, especially IMF<sub>1</sub>, most of the information and complexity from the original data are retained, making it a modal component with high information content.

Therefore, IMF<sub>1</sub> maintains a relatively high level of complexity in describing molten iron temperature, and its permutation entropy is very close to that of the original data sequence.

This observation provides insights into the loss of information during the decomposition process and emphasizes the importance of IMF<sub>1</sub> in retaining data complexity and essential features. IMF<sub>1</sub> retains the same level of complexity as the original data while preserving critical information from the original dataset. Therefore, through visual observation of the decomposed sequences, correlation analysis and permutation entropy analysis, this study chooses to use IMF<sub>1</sub> for regression modeling rather than employing other parameters for time series modeling. The primary consideration is that sequence one contains unique characteristics and causal relationships related to molten iron temperature compared to other sequences. Additionally, this decision simplifies the modeling process and reduces complexity, ensuring that the model is both effective and efficient in capturing the dynamics of molten iron temperature.

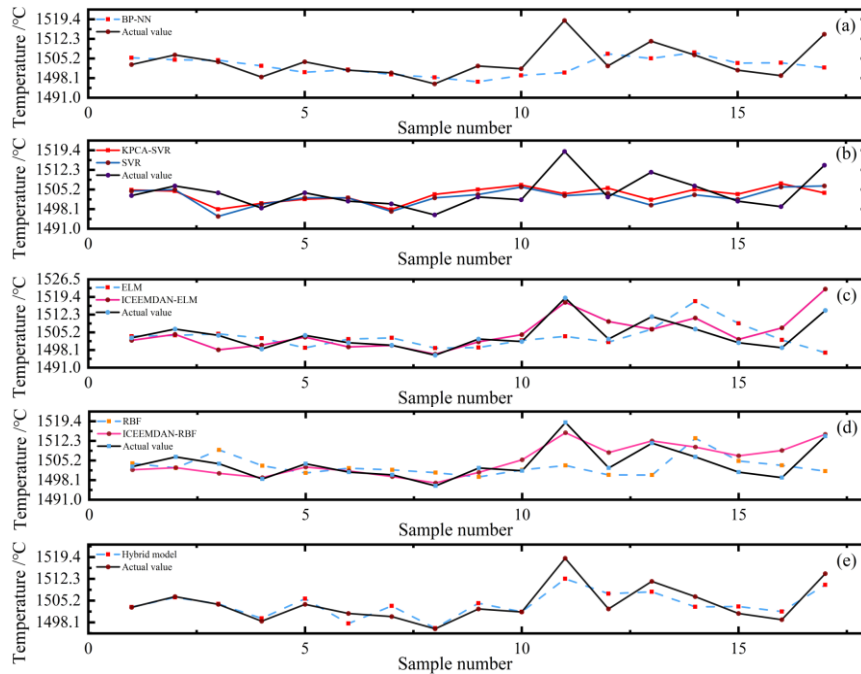
### 3.3. Model prediction and comparison

The temperature data of blast furnace hot metal after ICEEMDAN decomposition is used for a specific purpose. IMF<sub>1</sub> and related parameters are dimensionally reduced and then subjected to K-means clustering, dividing the data into two clusters. These two clusters are then separately input into SVR for training and testing. The training and testing data sets are split in an 8:2 ratio, with 63 data points used for training and 17 data points for testing. The training and testing of RBF-NN are also performed. The remaining IMFs and residual are used for further exploration. In this setup, each sample is input into an RBF-NN with a sample length of 4, meaning it uses the first four temperature data points to predict the following temperature data point. The parameters of the eight model are set as shown in Table 4.

**Table 4.** Training parameters in terms of the various existing and proposed models.

Model	Related training parameters
BP-NN	Iteration count is 1000 Error threshold is $1 \times 10^{-6}$ Learning rate is 0.01
SVR	Penalty factor is $c=4.0$ Radial basis function parameter is $g=0.8$
KPCA-SVR	Penalty factor is $c=4.0$ Radial basis function parameter is $g=0.8$
ELM	Number of hidden layer nodes is 20 Activation function is sigmoid
ICEEMDAN-ELM	Number of hidden layer nodes is 20 Activation function is sigmoid
RBF-NN	Radial basis function spread is 8000
ICEEMDAN-RBF-NN	Radial basis function spread is 8000
K-means	$K=2$
Hybrid model	SVR Penalty factor is $c=4.0$ Radial basis function parameter $g=0.8$
	RBF-NN Radial basis function spread is 8000

The prediction results are depicted in Figure 6. As shown in this figure, the proposed hybrid model can accurately forecast temperature fluctuations. The alignment between the model output curve and the actual data curve is quite high. This approach is beneficial for effective advanced temperature control in the furnace.



**Figure 6.** Results of the hybrid model and temperature predictions of the other seven models for molten iron temperature.

To thoroughly validate the performance of the proposed model, a comparative study was conducted with four classical prediction methods. The compared models include BP-NN, SVR, RBF-NN and ELM models. Additionally, comparative experiments were conducted using ICEEMDAN-ELM and ICEEMDAN-RBF-NN models to assess the improvements brought by the ICEEMDAN algorithm in model performance. To quantitatively evaluate the model performance, several metrics were used. In fact, MAPE measures the average deviation of the forecasted results from the actual results, RMSE represents the degree of deviation between model predictions and measurement reference values, PCC characterizes the correlation between the prediction curve and the actual curve and HR tests the accuracy of the model predictions. The results of these evaluations are presented in Table 4.

As indicated in Table 5: (1) from the perspective of prediction accuracy and reliability, the composite model proposed in this paper achieves a MAPE of 0.15 and an RMSE of 2.94, both significantly lower than those of traditional models. Specifically, the MAPE and RMSE are reduced by 54.55% and 49.40% compared to BP-NN, 50.00% and 53.18% compared to SVR, 54.55% and 56.83% compared to ELM and 46.43% and 53.33% compared to RBF-NN. This demonstrates a substantial improvement in prediction accuracy and reliability with the new composite model. (2) Considering the overall performance of the model, the proposed model achieves a PCC of 0.87, indicating a strong correlation between the predicted results and the actual values. In contrast, traditional models exhibit only a weak correlation (i.e., 0.44, 0.73, 0.77 and 0.72 higher than traditional models). The proposed model shows a stronger ability to capture underlying patterns. (3) From the



perspective of hit rates, the proposed model achieves hit rates as high as 94.12% when allowing an error margin of  $\pm 5$ . This represents a significant improvement over traditional models (i.e., an increase of 41.18%, 29.41%, 23.53% and 17.65% compared to traditional models). These results demonstrate that within an acceptable error range, the proposed model consistently performs well and is more reliable. (4) Compared to single models, the use of the ICEEMDAN algorithm significantly enhances the prediction accuracy and hit rate of the composite model. This enhancement is particularly noticeable in the PCC metric, where models employing the ICEEMDAN algorithm outperform single models. ICEEMDAN-ELM exhibits a 0.69 increase in PCC compared to the ELM model, and ICEEMDAN-RBF-NN shows a 0.63 increase in PCC compared to the RBF-NN model. This suggests that the use of decomposition algorithms can improve the model prediction stability. (5) Furthermore, compared to ICEEMDAN-ELM and ICEEMDAN-RBF-NN models that use decomposition algorithms, the combination of K-means and SVR to enhance the first decomposed sequence (i.e., IMF<sub>1</sub>) has a significant impact. Hit rates are increased by 23.23% and 11.77%, respectively. It is indicated that the composite model proposed here is better suited for iron water temperature prediction, as evidenced by the lower MAE and RMSE values and the PCC value close to 1, demonstrating its excellence in accuracy, stability and interpretability. (6) In comparison to ICEEMDAN-RBF-NN model, the proposed novel model exhibits improvements in all four indicators including MAPE, RMSE, PCC and HR (i.e., with increases of 16.67%, 21.81%, 11.54% and 14.29%, respectively). This indicates that after undergoing ICEEMDAN decomposition, the specific treatment of the IMF<sub>1</sub> sequence, which contains the original information and is the most complex, is crucial for enhancing accuracy.

**Table 5.** Evaluation indices of predictions obtained by the proposed model and other comparative models.

No.	Model	MAPE	RMSE	HR ( $\pm 5$ )	PCC
1	BP-NN	0.33	5.81	52.94%	0.43
2	RBF-NN	0.28	6.30	76.47%	0.15
3	ELM	0.33	6.81	70.59%	0.10
4	SVR	0.30	6.28	64.71%	0.14
5	KPCA-SVR	0.32	6.28	58.82%	0.03
6	ICEEMDAN-ELM	0.20	4.14	70.59%	0.79
7	ICEEMDAN-RBF-NN	0.18	3.76	82.35%	0.78
8	Hybrid model	0.15	2.94	94.12%	0.87

#### 4. Conclusions

To enhance the predictive accuracy of the blast furnace temperature forecasting model, a composite model incorporating ICEEMDAN, K-means, SVR and RBF-NN techniques for molten iron temperature prediction are proposed and validated. This new model amalgamates the strengths of ICEEMDAN, K-means, SVR and RBF-NN, enabling effective modeling and prediction of the intricate characteristics within the iron water temperature in the blast furnace smelting process. The findings reveal the following.

(1) The application of decomposition techniques facilitates the superior extraction of pertinent features from the temperature data. This, in turn, enhances prediction accuracy and provides a more

precise and interpretable decomposition of the iron water temperature sequence.

(2) In comparison to traditional models including BP-NN, SVR and ELM, the new model exhibits a higher degree of precision in forecasting iron water temperature, thereby increasing the reliability and stability of the forecasted temperature.

(3) Simultaneously exploring the relationships between relevant factors and time series enables a deeper understanding of the complexities involved in the blast furnace smelting process. This in-depth investigation contributes valuable insights to the prediction model.

(4) Within the ICEEMDAN decomposition of furnace temperature time series, the first decomposed sequence, namely  $IMF_1$ , demonstrates similarities in terms of correlation and sample entropy when compared to the rest of the sequences. It is suggested that  $IMF_1$  contains a substantial portion of the original sequence fluctuation information, indicating its significance and merit.

In summary, our findings in this paper demonstrate the effectiveness and improvements of the proposed hybrid model in predicting molten iron temperatures. Despite the current limited sample size of datasets of molten iron temperature, this work provides valuable insights for optimizing and controlling the blast furnace smelting process and holds significant potential for various applications. Future work could involve expanding the model application scope, and considering more real-world scenarios and factors to further enhance prediction accuracy and reliability. However, the need for interpretability is crucial for deploying intelligent algorithms as ironmaking processes undergo digital transformation. That aim is to enhance overall model understanding, improving its robustness and adaptability to diverse production environments. Therefore, future work would also focus on utilizing interpretable models or methods to offer a more intuitive understanding, aiming to guide the effective implementation of digitalization in iron and steel production. In that sense, upcoming indispensable efforts would specifically explore the impact of larger sample sizes on establishing the molten iron temperature model.

### Use of 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), Natural Science Foundation of Yunnan Province, China (No. 202101AU070031), Scientific and Technological Talent and Platform Project of Yunnan Province, China (No. 202305AO350011), Yibin Federation of Social Sciences Circles (No. 2023YBSKL101), Interdisciplinary Research Project of Kunming University of Science and Technology (No. KUST-xk2022001), Open Foundation of State Environmental Protection Key Laboratory of Mineral Metallurgical Resources Utilization and Pollution Control (No. HB202204) and Young Elite Scientist Sponsorship Program by China Association for Science and Technology, China (No. YESS20210106).

### Conflict of interest

The authors declare that they have no conflicts of interest.

## References

1. S. Liu, Q. Xiao, An empirical analysis on spatial correlation investigation of industrial carbon emissions using SNA-ICE model, *Energy*, **224** (2021), 120183. <https://doi.org/10.1016/j.energy.2021.120183>
2. K. He, L. Wang, A review of energy use and energy-efficient technologies for the iron and steel industry, *Renew. Sust. Energy Rev.*, **70** (2017), 1022–1039. <https://doi.org/10.1016/j.rser.2016.12.007>
3. M. A. Quader, S. Ahmed, R. A. R. Ghazilla, S. Ahmed, M. Dahari, A comprehensive review on energy efficient CO<sub>2</sub> breakthrough technologies for sustainable green iron and steel manufacturing, *Renew. Sust. Energy Rev.*, **50** (2015), 594–614. <https://doi.org/10.1016/j.rser.2015.05.026>
4. M. Smith, Blast furnace ironmaking: view on future developments, *Ironmak. Steelmak.*, **42** (2015), 734–742. <https://doi.org/10.1179/0301923315z.000000000422>
5. Y. Yang, S. Zhang, Y. Yin, A modified ELM algorithm for the prediction of silicon content in hot metal, *Neural Comput. Appl.*, **27** (2014), 241–247. <https://doi.org/10.1007/s00521-014-1775-x>
6. J. Li, X. Wei, H. Chen, Y. Yang, L. Min, Double-hyperplane fuzzy classifier design for tendency prediction of silicon content in molten iron, *Fuzzy Set Syst.*, **426** (2022), 163–175. <https://doi.org/10.1016/j.fss.2021.05.002>
7. R. D. Martín, F. Obeso, J. Mochón, R. Barea, J. Jiménez, Hot metal temperature prediction in blast furnace using advanced model based on fuzzy logic tools, *Ironmak. Steelmak.*, **34** (2013), 241–247. <https://doi.org/10.1179/174328107x155358>
8. C. S. Tsao, R. H. Day, A process analysis model of the U.S. steel industry, *Manage. Sci.*, **17** (1971), B569–B704. <https://doi.org/10.1287/mnsc.17.10.B588>
9. Y. Dai, J. Li, C. Shi, W. Yan, Dephosphorization of high silicon hot metal based on double slag converter steelmaking technology, *Ironmak. Steelmak.*, **48** (2020), 447–456. <https://doi.org/10.1080/03019233.2020.1807288>
10. J. Y. Liu, W. Zhang, Blast furnace temperature prediction based on RBF neural network and genetic algorithm, *Electron. Meas. Technol.*, **41** (2018), 42–45. <https://doi.org/10.3390/lubricants9090086>
11. V. R. Radhakrishnan, K. M. Ram, Mathematical model for predictive control of the bell-less top charging system of a blast furnace, *J. Process. Control*, **11** (2001), 565–586. [https://doi.org/10.1016/s0959-1524\(00\)00026-3](https://doi.org/10.1016/s0959-1524(00)00026-3)
12. M. Geerdes, R. Chaigneau, O. Linguardi, R. Molenaar, R. van Opbergen, Y. Sha, et al., *Modern blast furnace ironmaking*, IOS Press eBooks, 2020. <https://doi.org/10.3233/stal9781643681238>
13. S. Amano, T. Takarabe, T. Nakamori, H. Oda, M. Taira, S. Watanabe, et al., Expert system for blast furnace operation at Kimitsu works, *ISIJ Int.*, **30** (1990), 105–110. <https://doi.org/10.2355/isijinternational.30.105>
14. T. Yang, S. Yang, G. Zuo, H. Wei, J. Xu, Y. Zhou, An expert system for abnormal status diagnosis and operation guide of a blast furnace, *IFAC Proc. Vol.*, **25** (1992), 59–63. [https://doi.org/10.1016/S1474-6670\(17\)49899-5](https://doi.org/10.1016/S1474-6670(17)49899-5)
15. E. Lughofer, R. A. Pollak, C. Feilmayr, M. Schatzl, S. Saminger-Platz, Prediction and explanation models for hot metal temperature, silicon concentration, and cooling capacity in ironmaking blast furnaces, *Steel Res. Int.*, **92** (2021), 2100078. <https://doi.org/10.1002/srin.202100078>
16. J. Jimenez, J. Mochon, J. S. de Ayala, F. Obeso, Blast furnace hot metal temperature prediction through neural networks-based models, *ISIJ Int.*, **44** (2004), 573–580. <https://doi.org/10.2355/isijinternational.44.573>

17. W. Chen, F. Kong, B. Wang, Y. H. Li, Application of grey relational analysis and extreme learning machine method for predicting silicon content of molten iron in blast furnace, *Ironmak. Steelmak.*, **46** (2018), 974–979. <https://doi.org/10.1080/03019233.2018.1470146>
18. J. Song, X. Xing, Z. Pang, M. Lv, Prediction of silicon content in the hot metal of a blast furnace based on FPA-BP model, *Metals*, **13** (2023), 918. <https://doi.org/10.3390/met13050918>
19. W. Liang, G. Wang, X. Ning, J. Zhang, Y. Li, C. Jiang, et al., Application of BP neural network to the prediction of coal ash melting characteristic temperature, *Fuel*, **260** (2020), 116324. <https://doi.org/10.1016/j.fuel.2019.116324>
20. X. Su, S. Zhang, Y. Yin, W. Xiao, Prediction model of hot metal temperature for blast furnace based on improved multi-layer extreme learning machine, *Int. J. Mach. Learn. Cyb.*, **10** (2019), 2739–2752. <https://doi.org/10.1007/s13042-018-0897-3>
21. H. Zhang, Y. Yin, S. Zhang, An improved ELM algorithm for the measurement of hot metal temperature in blast furnace, *Neurocomputing*, **174** (2015), 232–237. <https://doi.org/10.1016/j.neucom.2015.04.106>
22. X. Huang, H. Chen, X. Ling, L. Liu, T. Huhe, Investigation of heat and mass transfer and gas-liquid thermodynamic process paths in a humidifier, *Energy*, **261** (2022), 125156. <https://doi.org/10.1016/j.energy.2022.125156>
23. M. M. Li, B. Verma, Nonlinear curve fitting to stopping power data using RBF neural networks, *Expert Syst. Appl.*, **45** (2016), 161–171. <https://doi.org/10.1016/j.eswa.2015.09.033>
24. K. Yang, J. Li, M. Wang, H. Wang, Q. Xiao, Identifying flow patterns in a narrow channel via feature extraction of conductivity measurements with a support vector machine, *Sensors*, **23** (2023), 1907. <https://doi.org/10.3390/s23041907>
25. P. Zhou, D. Guo, H. Wang, T. Chai, Data-driven robust M-LS-SVR-based NARX modeling for estimation and control of molten iron quality indices in blast furnace ironmaking, *IEEE Trans. Neural Networks Learn. Syst.*, **29** (2018), 4007–4021. <https://doi.org/10.1109/tnnls.2017.2749412>
26. P. Zhou, D. Guo, T. Chai, Data-driven predictive control of molten iron quality in blast furnace ironmaking using multi-output LS-SVR based inverse system identification, *Neurocomputing*, **308** (2018), 101–110. <https://doi.org/10.1016/j.neucom.2018.04.060>
27. Y. Liao, Y. Wang, M. Li, Q. Xiao, H. Wang, Prediction for the reduction smelting temperature based on CEEMDAN-RVM-EC, *Control Theory Appl.*, **39** (2022), 2177–2184.
28. Y. Wang, P. Yang, S. Zhao, J. Chevallier, Q. Xiao, A hybrid intelligent framework for forecasting short-term hourly wind speed based on machine learning, *Expert. Syst. Appl.*, **213** (2023), 119223. <https://doi.org/10.1016/j.eswa.2022.119223>
29. K. Yang, Y. Wang, M. Li, X. Li, H. Wang, Q. Xiao, Modeling topological nature of gas-liquid mixing process inside rectangular channel using RBF combined with CEEMDAN-VMD, *Chem. Eng. Sci.*, **267** (2023), 118353. <https://doi.org/10.1016/j.ces.2022.118353>
30. Y. Wang, P. Yang, Z. Song, J. Chevallier, Q. Xiao, Intelligent prediction of annual CO<sub>2</sub> emissions under data decomposition mode, *Comput. Econ.*, 2023. <https://doi.org/10.1007/s10614-023-10357-8>
31. Z. Cui, A. Yang, L. Wang, Y. Han, Dynamic prediction model of silicon content in molten iron based on comprehensive characterization of furnace temperature, *Metals*, **12** (2022), 1403. <https://doi.org/10.3390/met12091403>
32. D. O. L. Fontes, L. G. S. Vasconcelos, R. P. Brito, Blast furnace hot metal temperature and silicon content prediction using soft sensor based on fuzzy C-means and exogenous nonlinear autoregressive models, *Comput. Chem. Eng.*, **141** (2020), 107028. <https://doi.org/10.1016/j.compchemeng.2020.107028>

33. J. Zhao, X. Li, L. Song, K. Wang, Q. Lyu, E. Liu, Prediction of hot metal temperature based on data mining, *High Temp. Mat. Process.*, **40** (2021), 87–98. <https://doi.org/10.1515/htmp-2021-0020>
34. L. J. Cao, K. S. Chua, W. Chong, H. P. Lee, Q. M. Gu, A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine, *Neurocomputing*, **55** (2003), 321–336. [https://doi.org/10.1016/s0925-2312\(03\)00433-8](https://doi.org/10.1016/s0925-2312(03)00433-8)
35. K. Yang, H. Wang, H. Wang, M. Li, Q. Xiao, Topological approach for the measurement of mixing state quality in a vertical rectangular channel, *Int. J. Multiphas. Flow*, **163** (2023), 104431. <https://doi.org/10.1016/j.ijmultiphaseflow.2023.104431>



AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)