processes

MDPI

Article

Predicting Alloying Element Yield in Converter Steelmaking Using t-SNE-WOA-LSTM

Xin Liu [1], Xihui Qu [2], Xinjun Xie [2], Sijun Li [2], Yanping Bao [1,*] and Lihua Zhao [3,*]

[1] State Key Laboratory of Advanced Metallurgy, University of Science & Technology Beijing, Beijing 100083, China
[2] Shan Dong Iron and Steel Co., Ltd., Jinan 271104, China
[3] Metallurgical and Ecological Engineering School, University of Science & Technology Beijing, Beijing 100083, China
* Correspondence: baoyp@ustb.edu.cn (Y.B.); zhaolihua@metall.ustb.edu.cn (L.Z.)

Abstract: The performance and quality of steel products are significantly impacted by the alloying element control. The efficiency of alloy utilization in the steelmaking process was directly related to element yield. This study analyses the factors that influence the yield of elements in the steelmaking process using correlation analysis. A yield prediction model was developed using a t-distributed stochastic neighbor embedding (t-SNE) algorithm, a whale optimization algorithm (WOA), and a long short-term memory (LSTM) neural network. The t-SNE algorithm was used to reduce the dimensionality of the original data, while the WOA optimization algorithm was employed to optimize the hyperparameters of the LSTM neural network. The t-SNE-WOA-LSTM model accurately predicted the yield of Mn and Si elements with hit rates of 71.67%, 96.67%, and 99.17% and 57.50%, 89.17%, and 97.50%, respectively, falling within the error range of ±1%, ±2%, and ±3% for Mn and ±1%, ±3%, and ±5% for Si. The results demonstrate that the t-SNE-WOA-LSTM model outperforms the backpropagation (BP), LSTM, and WOA-LSTM models in terms of prediction accuracy. The model was applied to actual production in a Chinese plant. The actual performance of the industrial application is within a ±3% error range, with an accuracy of 100%. Furthermore, the elemental yield predicted by the model and then added the ferroalloys resulted in a reduction in the elemental content of the product by 0.017%. The model enables accurate prediction of alloying element yields and was effectively applied in industrial production.

Keywords: alloy element yield; converter steelmaking; t-SNE; prediction model; industrial applications

Citation: Liu, X.; Qu, X.; Xie, X.; Li, S.; Bao, Y.; Zhao, L. Predicting Alloying Element Yield in Converter Steelmaking Using t-SNE-WOA-LSTM. *Processes* **2024**, *12*, 974. https://doi.org/10.3390/pr12050974

Academic Editor: Alexander S. Novikov

Received: 7 April 2024
Revised: 7 May 2024
Accepted: 8 May 2024
Published: 10 May 2024

1. Introduction

The converter steelmaking process is the primary method of long-flow steelmaking in China's iron and steel industry, resulting in significant carbon emissions [1,2]. The steel industry must address the urgent issue of reducing energy consumption and carbon emissions, given the implementation of the Peak Carbon strategy [3,4]. Ferroalloys are essential raw materials in converter steelmaking used to regulate steel composition and achieve standard mechanical properties, accounting for 5–10% of the cost of steelmaking [5]. The quantity of the alloy in the steel has a significant impact on its quality. An excessively high amount will result in a waste of resources, which will ultimately lead to the scrapping of steel. Conversely, a deficiency in the alloy will result in the product failing to meet the requisite performance standards and increasing production costs. The yield of the alloying element is a crucial parameter for calculating the quantity of the alloy and represents the primary indicator for evaluating the utilization efficiency of ferroalloys. Currently, the majority of enterprises rely on manual experience to determine alloy yield, which often results in significant discrepancies. The accurate determination of alloy element yield is conducive to the 'narrow control' of the liquid steel. This ultimately achieves the

objective of energy saving and emission reduction. Consequently, the accurate prediction of elemental yield is of paramount importance in converter steelmaking.

The conventional approach to predicting elemental yield involves using historical data. The average yield from past furnaces is taken as the current furnace's elemental yield and then used to calculate the necessary alloy addition [6]. If the historical furnace data fluctuate or do not match the current furnace data, the predicted result will have a significant error compared to the actual value. Due to the improvement of steelmaking equipment, Artificial Intelligence (AI) has become increasingly important in modern steelmaking [7]. The implementation of AI algorithms in the steelmaking process can achieve precise prediction and control of the process [8], optimal control of material consumption [9], analysis of working conditions, and quality assessment [10]. Neural network algorithms are characterized by easy feature extraction, high generalization, and high model adaptability [11]. The steelmaking process involves complex physicochemical reactions and numerous operating means. The neural network possesses remarkable adaptability and intelligence. It can adjust its network structure and function according to different application scenarios, demonstrating its flexibility and versatility [12].

Calculating elemental yields using neural network algorithms is one of the most effective ways to improve prediction accuracy. Zheng et al. developed a combined model using particle swarm optimization (PSO) and long short-term memory (LSTM). This model was specifically designed to support the HRB400E production process. After applying this model, the efficiency of steel composition control increased by 4% [13]. The LSTM model has a relatively slow training process due to its complex gating mechanism and multiple weighting parameters [14]. Xu et al. developed a support vector regression (SVR) model to predict the elemental yield of the LF furnace [15]. While the prediction accuracy is high, the training time for the SVR model is considerable. This is particularly true when dealing with large amounts of data. The computational complexity associated with such large datasets can become a significant challenge for the stable operation of the SVR model. The choice of hyperparameters has a significant impact on the performance of the SVR model, which is prone to overfitting or underfitting [16]. Table 1 compares the prediction models designed by the researchers for the alloying element yield. The IPSO-ELM model obtains a 95% hit rate, is fast to train, and is suitable for solving high-dimensional problems but is affected by random initial weights. The ENN model has a strong learning ability and memory function, but the computational complexity is relatively high. The Mn element yield prediction models all achieved an accuracy of over 84.50%, with the PCA-DNN model achieving the highest accuracy at 99.50% for predicting Mn with an error of ±3%. However, the high complexity of DNN models makes training difficult and model interpretability poor. It is time consuming in industrial application engineering and is not suitable for the fast-paced production of converter steelmaking. Thus, it is necessary to create a prediction model that has high computational accuracy and data processing capabilities to achieve precise control of alloying elements.

Table 1. Comparison of prediction models for the converter endpoint.

Alloying Element	Model	Hit Rate	Advantages	Disadvantages	Reference
Mn	MLR	84.50%	(1) Non-linear learning; (2) scalability; (3) sparsity	(1) Multicollinearity problem; (2) high correlation among total variables	[17]
Mn	PCA–DNN	99.50%	(1) Highly non-linear problems; (2) large-scale data	(1) Difficult to train; (2) poor model interpretability	[6]

Table 1. *Cont.*

Alloying Element	Model	Hit Rate	Advantages	Disadvantages	Reference
Mn	AO-ENN	86.00%	(1) Stronger learning and memory skills; (2) easy to train; (3) highly interpretable	(1) Relatively high computational complexity; (2) easy to fall into local minima	[18]
Mn	GA-BP	98.42%	(1) Strong non-linear mapping capability; (2) high self-learning and self-adaptive capabilities; (3) some fault tolerance	(1) Slow convergence speed; (2) easy to fall into local minima	[19]
Mn	SVM	88.57%	(1) Generalization properties; (2) high-dimensional issues; (3) avoiding neural network structure selection and local minima problems	(1) Sensitive to missing data; (2) no generic solution to non-linear problems	[20]
Mn	IPSO-ELM	95.00%	(1) Fast training; (2) high-dimensional issues; (4) avoiding neural network structure selection and local minima problems	(1) Poor model interpretability; (2) affected by random initialization weights	[21]
Si	MLR	73.50%	(1) Non-linear learning; (2) scalability; (3) sparsity	(1) Multicollinearity problem; (2) high correlation among total variables	[17]
Si	PCA–DNN	98.80%	(1) Highly non-linear problems; (2) large-scale data	(1) Difficult to train; (2) poor model interpretability	[6]
C	AO-ENN	88.00%	(1) Stronger learning and memory skills; (2) easy to train; (3) highly interpretable	(1) Relatively high computational complexity; (2) easy to fall into local minima	[18]

This paper proposes a prediction model for the Mn/Si yield of the Q345B smelting process. The model combines the t-distributed stochastic neighbor embedding (t-SNE) algorithm with the long short-term memory (LSTM). Additionally, the whale optimization algorithm (WOA) is utilized to enhance the model's performance. The t-SNE algorithm is used to reduce the dimensionality of production process data. This allows for the expression of the internal structural relationships of high-dimensional data in a lower-dimensional space [22]. This is suitable for the converter steelmaking process, which involves complex reactions and a large amount of data. The elemental yield of the steelmaking process is presented as time series data, and the LSTM neural network is chosen to establish the model. To achieve fast and accurate hyperparameter optimization of the LSTM neural network, introduce the WOA optimization algorithm, which addresses the difficulty of optimizing hyperparameters for LSTM neural networks. Comparing the prediction results of the established model with those of the backpropagation (BP), LSTM, and WOA-LSTM models and applying them to industrial production, it was found that the Mn/Si yield prediction model based on t-SNE-WOA-LSTM established in this paper has high prediction accuracy and good generalization performance. The model accurately predicts Mn/Si yield and is effective for industrial production.

2. Data Preprocessing and Methodology

2.1. Data Preprocessing

The data were obtained from a steel company located in eastern China, with Q345B serving as the subject of the research. The composition of Q345B steel is shown in Table 2. From October to December 2023, 971 data items were collected to record converter and refining production process data and molten steel composition. The data included 16 variable indicators. The following measures were applied to the raw data for data processing:

Table 2. Composition of Q345B steel (wt%).

C	Si	Mn	P	S
0.15–0.17	0.41–0.45	1.42–1.46	≤0.025	≤0.010

(1) Due to the large amount of data and the small number of missing values, the furnace containing missing values is rejected. Firstly, the data rows containing missing values are filtered out and the IFERROR function is selected for elimination;

(2) The abnormal zero values of certain variables are filled with mean value processing. Screen out data lines containing abnormal zero values and select the AVERAGE function to fill;

(3) In order to reduce the impact of repeated values on the prediction accuracy, the repeated item parameters in the dataset are removed;

(4) The outliers are removed using the 3-fold mean square deviation method. Firstly, the standard deviation of the sample is calculated, then, the permissible range of the sample is determined, and finally, extreme outliers are removed.

(5) The data were normalized using Equation (1).

Experiments were conducted using a total of 800 sets of data after screening. The maximum, minimum, average, and range of values for each indicator variable are counted in Table 3.

$$x_i^* = \frac{x_i - x_{i\ min}}{x_{i\ max} - x_{i\ min}} \quad (1)$$

where x_i is the input feature variable and $x_{i\ max}$ and $x_{i\ min}$ are the maximum and minimum values of each independent sample data.

Table 3. Statistics of main parameters of indicator variables.

No	Indicator Variables	Maximum	Minimum	Mean	Range
X_1	Molten iron loading quantity [t]	88.12	70.08	80.01	18.04
X_2	Scrap steel loading quantity [t]	13.8	0	6.50	13.8
X_3	Number of turndown/times	3	1	1.43	2
X_4	Tapping temperature [°C]	1682	1535	1602.87	147
X_5	Nitrogen consumption [m^3]	1734	306	956.74	1428
X_6	Oxygen consumption [m^3]	3962	1911	2668.16	2051
X_7	Lime addition amount [kg]	3361	591	1488.40	2770
X_8	Light-burned dolomite addition amount [kg]	2592	0	823.72	2592
X_9	Tapping weight [t]	91.6	72	82.30	19.6
X_{10}	C content of molten iron [%]	0.14	0.02	0.066	0.12
X_{11}	Mn content of molten iron [%]	0.158	0.005	0.046	0.153
X_{12}	Ferrosilicon alloy [kg]	190	0	73.1	190
X_{13}	Silicomanganese alloy [kg]	2045	881	1589.53	1164
X_{14}	High carbon manganese alloy [kg]	403	0	40.90	403
X_{15}	Medium-carbon manganese alloy [kg]	512	0	62.99	512
X_{16}	Low-carbon manganese alloy [kg]	599	0	24.51	599
Y_1	Yield of the Si element [%]	92.84	66.34	77.43	26.58
Y_2	Yield of the Mn element [%]	99.28	80.25	89.31	18.33

The calculation formula of the alloying element yield is shown in Equation (2) [6] as follows:

$$\eta_i = \frac{[w_{ia} - w_{ib}] \times G}{w_i \times G_k} \times 100\% \quad (2)$$

where η_i is the yield of element i in alloy k; w_i is the content of element i in alloy k; G_k is the weight of alloy k, kg; G is the weight of molten steel, kg; w_{ia} is the final content of element i after alloying; and w_{ib} is the initial content of element i before alloying.

2.2. Methodology of t-SNE-WOA-LSTM

t-SNE is a non-linear method for reducing data dimensionality. It has the advantage of identifying data manifold features using the stream shape learning method and projecting them onto a low-dimensional space while preserving local information [23,24]. The t-SNE algorithm converts the high-dimensional Euclidean distance between data points into a conditional probability that represents similarity. The conditional probability of similarity $p_{j|i}$ of data point x_i with respect to data point x_j can be expressed mathematically.

$$p_{j|i} = \frac{\exp\left(-\|x_i - x_j\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq l} \exp\left(-\|x_i - x_k\|^2 / 2\sigma_i^2\right)} \quad (3)$$

where σ_i^2 is the Gaussian variance centered on the data point x_j. The pairwise similarity $p_{j|i}$ value is 0.

For the low-dimensional representations y_i and y_j of data point x_i and x_j, their conditional similarity probabilities $q_{j|i}$ can be computed accordingly in mathematical form as follows:

$$q_{j|i} = \frac{\exp\left(-\|y_i - y_j\|^2\right)}{\sum_{k \neq l} \exp\left(-\|y_i - y_k\|^2\right)} \quad (4)$$

Since y_i and y_j model the state of data point x_i and data point x_j in the low-dimensional space, the conditional probability of similarity between two points in these two spaces is equal, i.e., $p_{j|i} = q_{j|i}$.

The t-SNE algorithm achieves dimensionality reduction by minimizing the error between $p_{j|i}$ and $q_{j|i}$. The loss function C is computed by minimizing the KL (Kullback–Leibler) scatter of $p_{j|i}$ and $q_{j|i}$ through a gradient descent algorithm as follows:

$$C = \sum_i KL(P_i \| Q_i) \quad (5)$$

where P_i denotes the conditional probability distribution of all other data points on a given data point x_i and Q_i denotes the conditional probability distribution of all other data points on a given data point y_i. The final convergence result is obtained by minimizing the loss function C.

LSTM neural networks are excellent variants of the recurrent neural network (RNN) [25,26]. LSTM neural networks were developed to address the limitations of conventional RNNs, such as vanishing gradients and the inability to capture long-time dependencies in sequences. Figure 1 illustrates the neural network structure. The LSTM model creates a distinct pathway for transmitting long-term dependencies. Each LSTM unit performs basic addition and subtraction operations on this pathway to preserve the long-term dependency information. The training and inference time for LSTM neural networks is long due to high computational complexity. To mitigate this, the WOA optimization algorithm optimal exploration mechanism is used to quickly find the global optimal solution with high convergence efficiency and computational speed [27].

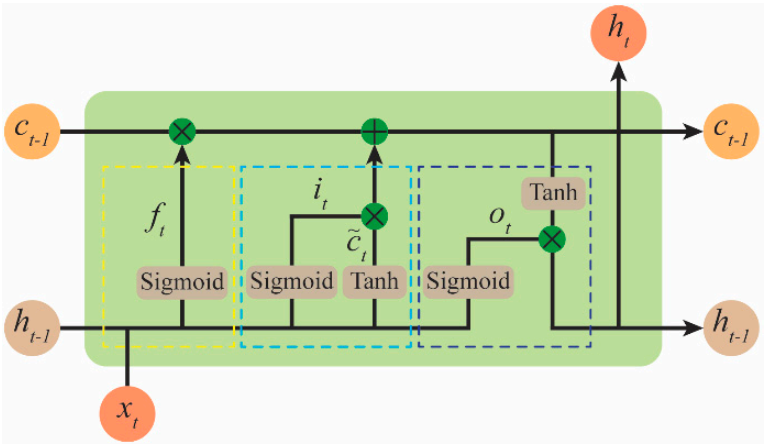

Figure 1. The structure of long short-term memory [28].

The WOA is an algorithm that simulates the hunting behavior of whales rounding up prey using random or optimal search agents and spirals to simulate the bubble net attack mechanism of humpback whales. The WOA is characterized by its simple mechanism, few parameters, and strong optimization capability [29]. The distance and position vectors between individuals in the WOA are as follows:

$$\begin{cases} D = \left| CX^*_{(t)} - X_{(t)} \right| \\ X_{(t+1)} = X^*_{(t)} - AD \end{cases} \tag{6}$$

where t denotes the current number of iterations, $X_{(t)}$ vector is the position vector, $X^*_{(t)}$ is the position vector of the best solution obtained so far, and A and C are the coefficient vectors: $A = 2a \times r_1 = a$, $C = 2 \times r_2$. During the iteration, the values of 2 and 0 should be decreased linearly. The variables r_1 and r_2 represent random vectors within the range of [0, 1].

The WOA employs two primary mechanisms for whale predation: encircling predation and bubble net predation. The position is updated based on the probability p of choosing either bubble net predation or contraction encirclement, as shown in Equation (7) as follows:

$$X_{(t+1)} = \begin{cases} X^*_{(t)} - AD & p \le 0.5 \\ D' e^{BL} \cos(2\pi L) + X^*_{(t)} & p > 0.5 \end{cases} \tag{7}$$

where D' is the distance between the current searching individual and the current optimal solution; Bis is the spiral shape parameter; and Lis is a uniformly distributed random number with value domain $[-1, 1]$. P is the probability of the predation mechanism, a random number with value domain [0, 1].

As the number of iterations t increases, the parameter A and the convergence factor gradually decrease. If $|A| < 1$, the whale swarm gradually surrounds the current optimal solution and is in the local optimization stage. To guarantee that all whales can search effectively in the solution space, the WOA updates the position based on the whales' distance from each other to achieve a random search. Therefore, when $|A| \ge 1$, the searching individual will swim towards the random whale to obtain the optimal solution.

$$\begin{cases} D'' = \left| CX_{rand(t)} - X_{(t)} \right| \\ X_{(t+1)} = X_{rand(t)} - AD \end{cases} \tag{8}$$

where D'' is the distance between the current search individual and the random individual. $X_{rand(t)}$ is the current position of the random individual.

Figure 2 shows the algorithmic flow of the elemental yield prediction model based on t-SNE-WOA-LSTM.

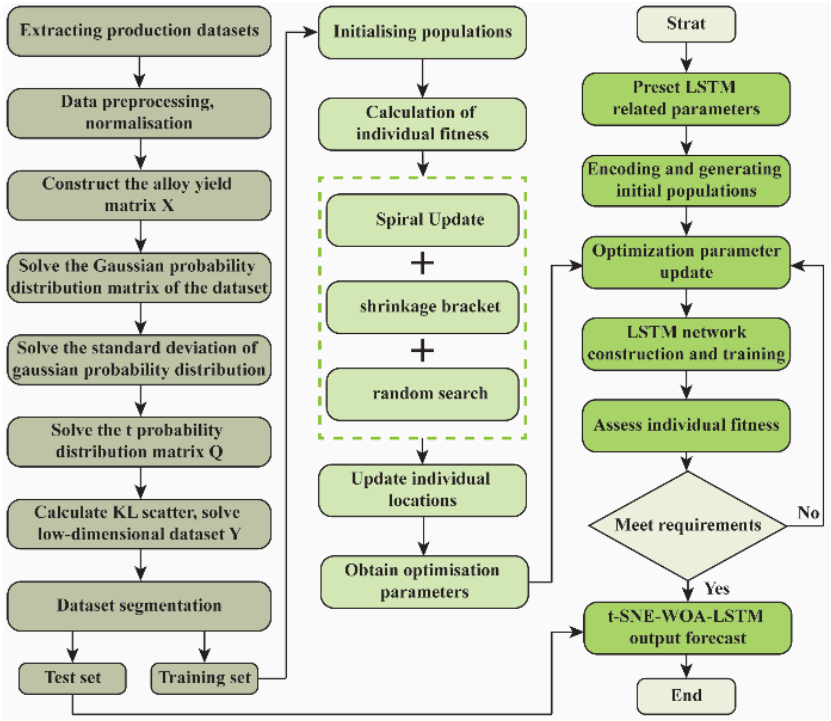

Figure 2. The flowchart of t-SNE-WOA-LSTM.

2.3. Evaluation Indicators

This study validates the algorithmic performance of the developed models using the same raw data. Model performance is evaluated using the root-mean-square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), coefficient of determination (R^2), and hit rates with different error ranges. The formulae for the indicators are as follows [30,31]:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - Y_i'\right)^2} \tag{9}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - Y_i'\right| \tag{10}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - Y_i'}{Y_i}\right| \times 100 \tag{11}$$

$$R^2 = \frac{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2 - \sum_{i=1}^{n}\left(Y_i - Y_i'\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2} \tag{12}$$

where Y_i, Y_i', and $\overline{Y}$ represent the actual, predicted, and average values, respectively, and n is the number of samples.

3. Results and Discussion

3.1. Correlation Analysis

The heat map of the Pearson correlation coefficient is shown in Figure 3. The correlation coefficient between the Mn content of molten iron and the silicomanganese alloy is −0.77, indicating a strong negative correlation. This means that the amount of the silicomanganese alloy added increases as the Mn content in the blow mold decreases. The correlation coefficient between the quantity of molten iron and the quantity of scrap loading

is −0.72, indicating a strong negative correlation. This means that as the quantity of iron loading increases, the quantity of scrap loading decreases. The correlation coefficient between the silicomanganese alloy and the high-carbon manganese alloy is −0.54, indicating a moderate negative correlation. Specifically, as the incorporation of the silicomanganese alloy increases, the incorporation of high-carbon ferromanganese decreases. The correlation coefficient between the number of turndowns and *C* content of molten iron is 0.50, which shows a more moderate positive correlation; the higher the *C* content of molten steel, the higher the number of turndowns.

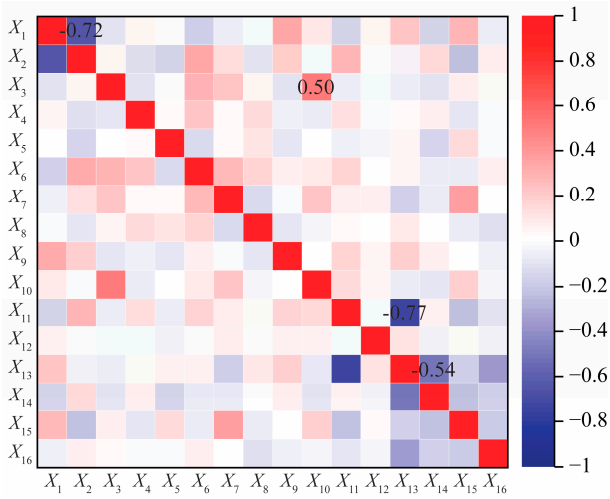

Figure 3. Heat map of the Pearson correlation coefficients.

3.2. The Effect of the Structure on the Prediction Results

The experiment selected the six network parameters of learning rate, dropout, batch size, iteration number, number of neurons, and time step using the control variable method. The RMSE and R^2 were used as the judgment criteria. In this section of the experiment, the processed dataset is used as the input variable, and the Mn/Si element yield is used as the only output variable to obtain the LSTM network parameters suitable for predicting the Mn/Si element yield.

The experiment controlled the dropout at 0.1, batch size at 30, iteration number at 3000, number of neurons at 50, and time step at 3. The LSTM neural network was built with an adjusted learning rate, and the resulting RMSE of the prediction model for Mn element yield is shown in Figure 4a. The optimal learning rate for the prediction model was found to be 0.01, as it produced the lowest RMSE and highest R^2. Therefore, a learning rate of 0.01 is recommended for the prediction model. Controlling the learning rate to 0.01, batch size to 30, number of iterations to 3000, number of neurons to 50, and time step to 3, the optimal parameter of dropout was obtained to be 0.05 (Figure 4b). In the same steps, the optimal hyperparameter batch size is 32 (Figure 4c), the number of iterations is 7000 (Figure 4d), the number of neurons is 20 (Figure 4e), and the time step is 4 (Figure 4f).

The optimal LSTM model parameters for Si element yield were experimentally determined using the same control variable method as shown in Figure 5. The hyperparameters of the LSTM model for Si yield prediction were optimized using the WOA to improve the poor fitting effect. As a result, the WOA-LSTM neural network model was obtained with a small error value. The WOA's initial population size is set to 50, with 35 iterations. The learning rate, dropout, batch size, number of iterations, and number of neurons are all within the range of [0.001,0.02], [0,0.1], [1,100], [5,1000], and [1,100], respectively. Figure 6 displays the optimization process for the adaptability of the element yield prediction model. It is evident that the adaptability value gradually decreases, indicating a reduction in the error between the predicted and actual values. The model fit improves progressively as the parameter combinations are optimized. The Mn element yields prediction model reached the optimal parameter combination at 18 iterations, while the Si element yields

prediction model reached the optimal parameter combination at 20 iterations. It indicates that the Si element yields prediction model requires more iteration time and a number of calculations in the hyperparameter optimization process. This is related to the influencing factors of the element Si in steelmaking, which is not only used in the alloying process but also deoxidation. Therefore, Si element yield data have more influence factors than Mn element yield, and the Si element yield prediction model hyperparametric optimization process requires a higher number of iterations. Table 4 presents the optimal parameter combinations for the Mn/Si element yield prediction model after 30 iterations.

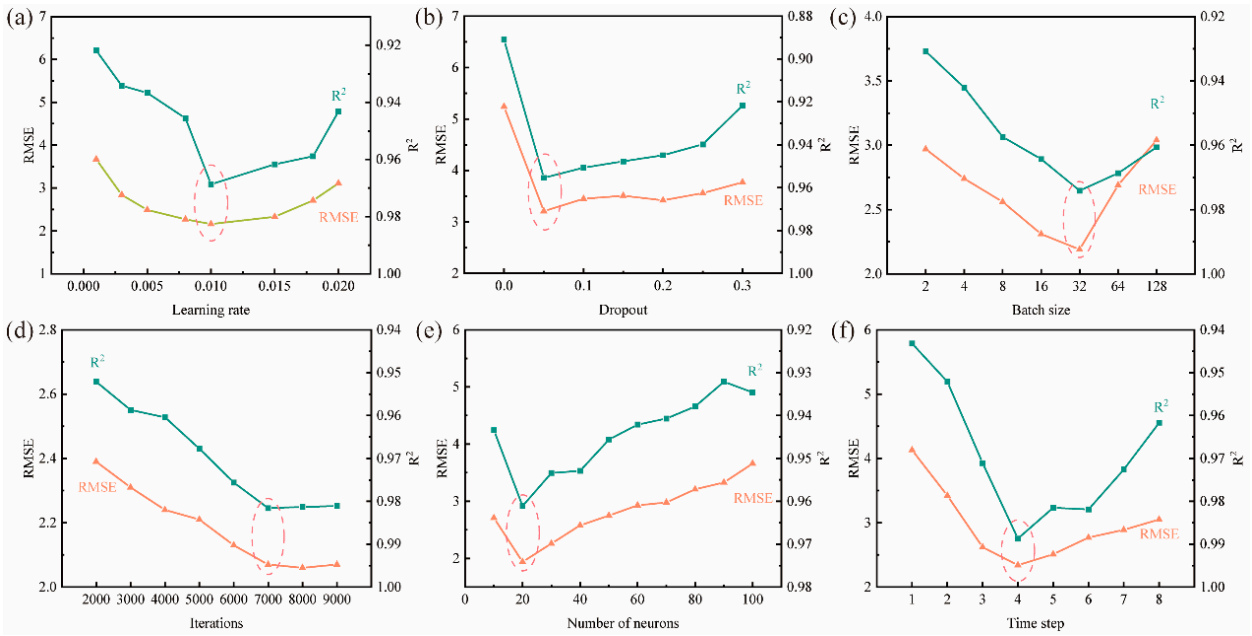

Figure 4. RMSE and R^2 with variations in (**a**) learning rate; (**b**) dropout rate; (**c**) batch size; (**d**) iterations; (**e**) number of neurons; and (**f**) time step for the LSTM prediction model of the Mn element yield.

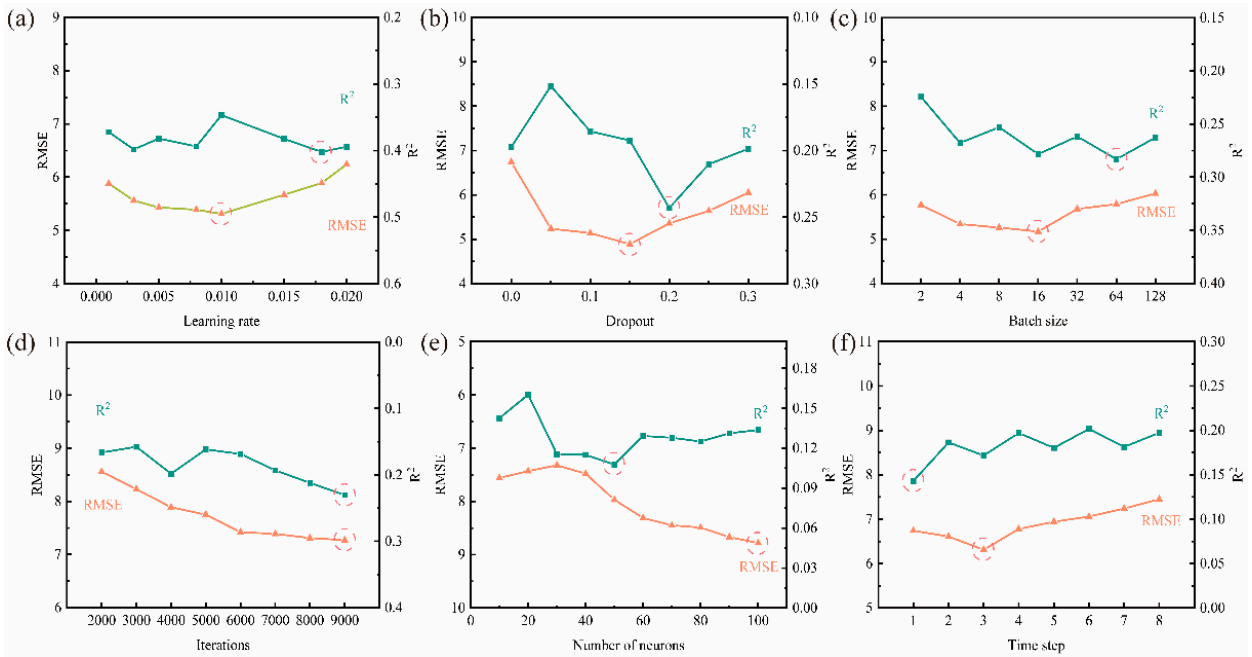

Figure 5. RMSE and R^2 with variations in (**a**) learning rate; (**b**) dropout rate; (**c**) batch size; (**d**) iterations; (**e**) number of neurons; and (**f**) time step for the LSTM prediction model of the Si element yield.

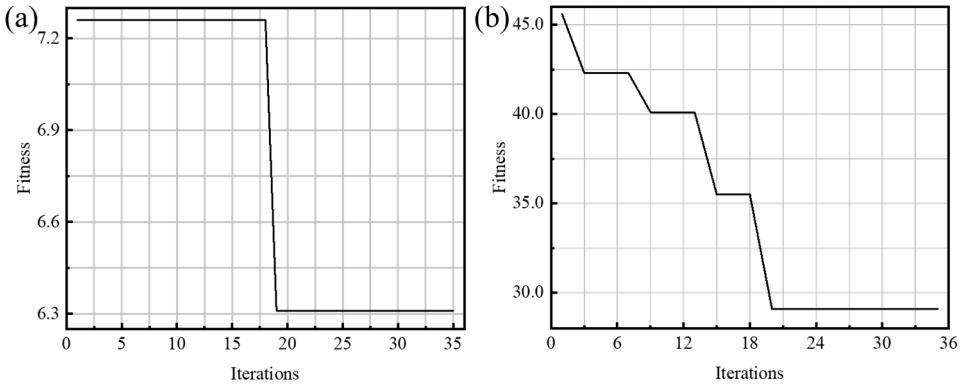

Figure 6. Parameter optimization of (**a**) Mn element yield; (**b**) Si element yield.

Table 4. Mn/Si element yield model optimal parameter combination.

Element	Learning Rate	Dropout Rate	Batch Size	Iterations	Number of Neurons	Time Step
Mn	0.0100	0.0500	32	194	30	4
Si	0.0043	0.4621	16	89	56	3

3.3. Comparison of the t-SNE-WOA–LSTM Model with Other Models

This study compares the commonly used BP neural network prediction model and the common LSTM neural network model with the t-SNE-WOA-LSTM model proposed in this study for predicting steel production data. The BP model uses a three-layer network structure with one hidden layer and five nodes in the hidden layer, with an initial learning rate of zero. The number of training iterations is set to 1000, with the activation function of neurons in the hidden layer selected as both tansig and purelin functions, and the training function chosen is the trainlm function.

Figure 7a displays the results of the BP neural network used to predict the yield of the Mn element. The RMSE is 2.065, the MAPE value is 1.871, and the R^2 is only 0.374. Therefore, the BP model has the worst prediction effect, and there is a large gap between the predicted value and the actual value. The prediction performance of the ordinary LSTM model is improved compared to BP, with an RMSE of 1.880, a MAPE value of 1.684, and an R^2 of only 0.535, as shown in Figure 7b. The traditional BP prediction model and the ordinary LSTM model are not effective in predicting data with slightly larger data volume and non-linear characteristics. For the WOA-LSTM neural network, their prediction accuracy is significantly improved compared to the previous two models, with the RMSE value decreasing by 0.733, the MAPE value decreasing by 0.729, and the R^2 value exceeding 0.7, as shown in Figure 7c. This indicates that the WOA-LSTM neural network model has a better fitting ability for data with non-linear characteristics and is able to process large amounts of data, which is a significant advantage in steelmaking yield prediction research. The t-SNE-WOA-LSTM model is further optimized based on the WOA-LSTM neural network, and the fitting ability is significantly improved, with the RMSE reaching 0.779, the MAPE reaching 0.571, and the R^2 close to 0.9, as shown in Figure 7d. Figure 9a shows the statistics of the hit rates of the four models in different ranges. The hit rate of the BP prediction model for Mn element yield is 38.33% within ±1% and 73.33% within ±3%. Under the condition of the same test dataset, the t-SNE-WOA-LSTM model has a hit rate of 71.67% within ±1% and 99.17% within ±3%, which is 33.34% and 25.84% higher, respectively, and the prediction accuracy is significantly improved.

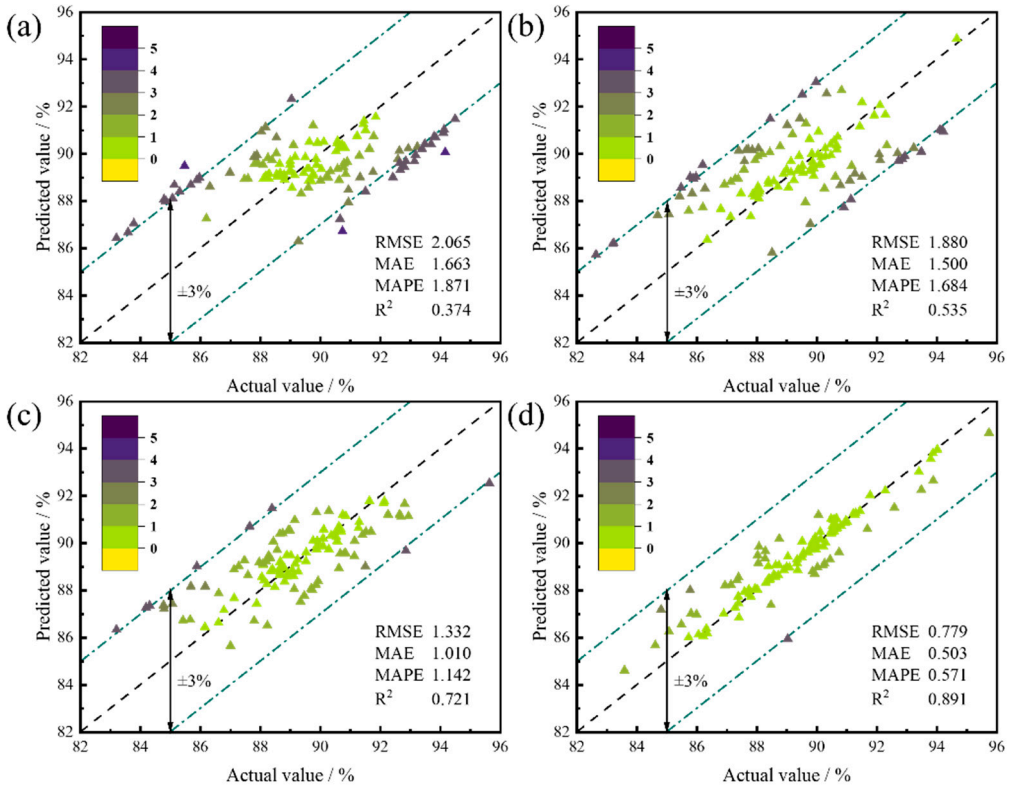

Figure 7. Comparison between actual and predicted Mn element yield values in (**a**) BP, (**b**) LSTM, (**c**) WOA–LSTM, and (**d**) the t-SNE-WOA–LSTM neural network prediction model.

Figure 8 shows the comparison of the effectiveness of the four neural networks in predicting Si element yield. It can be seen that the RMSE of the BP model is 3.840, the MAPE is 4.177, and the R^2 is 0.539. The accuracy of the BP model within ±1%, ±3%, and ±5% is 23.33%, 46.67%, and 69.17%, respectively. Therefore, the BP neural network model was not effective in predicting Mn/Si element yield, and the prediction error was large for some furnaces. The RMSE of the LSTM model is 3.551, the MAPE value is 3.630, and the R^2 is only 0.599. The LSTM model achieved hit rates of 30.83%, 53.33%, and 75.83% within ±1%, ±3%, and ±5%, respectively. Although the prediction effect is improved compared to the BP model, the fit is still low. The WOA-LSTM neural network has an RMSE of 2.551, a MAPE value of 2.407, and an R^2 of 0.730. The hit rate of the WOA-LSTM model within ±1%, ±3%, and ±5% is 44.17%, 75.83%, and 91.67%, respectively. The predicted values of Si elemental yield are closer to the actual values. This is because the WOA optimization algorithm autonomously optimizes the hyperparameters of the LSTM model, which speeds up prediction and improves accuracy. Figure 9b shows that the t-SNE-WOA-LSTM model predicts Si element yield with a high degree of accuracy. The hit rate within ±5% is 97.5%, which is significantly higher than the other three element yield prediction models.

A comparison was conducted between the computational results of this model and those of the BP, LSTM, and WOA-LSTM models. The t-SNE-WOA-LSTM prediction model was found to have the highest prediction accuracy. Comparison with the published models in Table 1 reveals that the prediction accuracy is higher than most models (Figure 10), with a 99.5% hit rate for the PCA-DNN model Mn prediction. PCA-DNN is utilized in the LF refining process, which has a long refining cycle, lower complexity compared to converter steelmaking, and lower computational requirements for the model. The t-SNE-WOA-LSTM model is well-suited to the converter steelmaking process, which exhibits complex response and a large data volume. It is highly efficient in terms of computing speed and adept at optimization searching, rendering it a more optimal choice for the actual production scenario of converter steelmaking than other models.

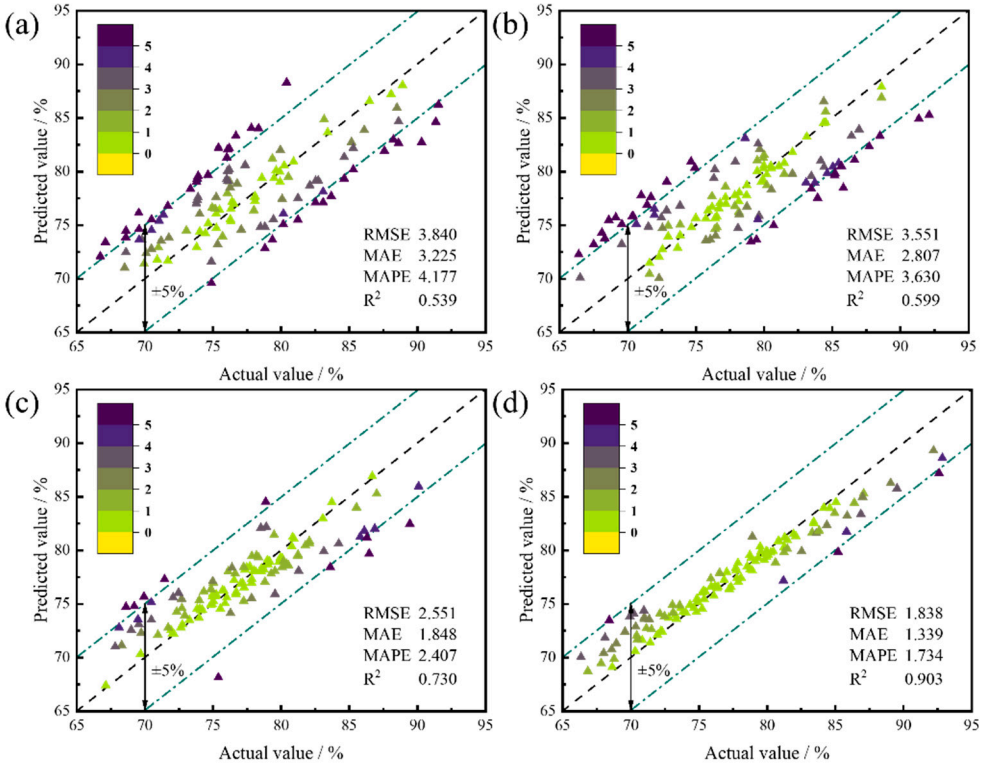

Figure 8. Comparison between actual and predicted Si element yield values in (**a**) BP, (**b**) LSTM, (**c**) WOA–LSTM, and (**d**) the t-SNE-WOA–LSTM neural network prediction model.

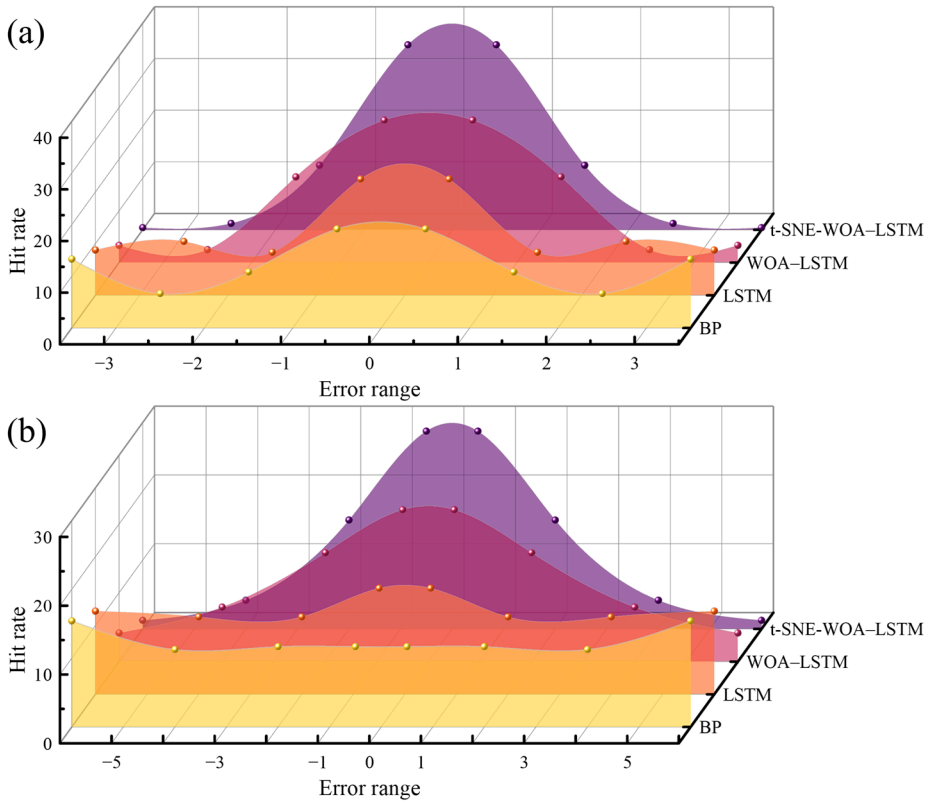

Figure 9. Hit rate predicted by (**a**) Mn and (**b**) Si element yield using different neural network prediction models.

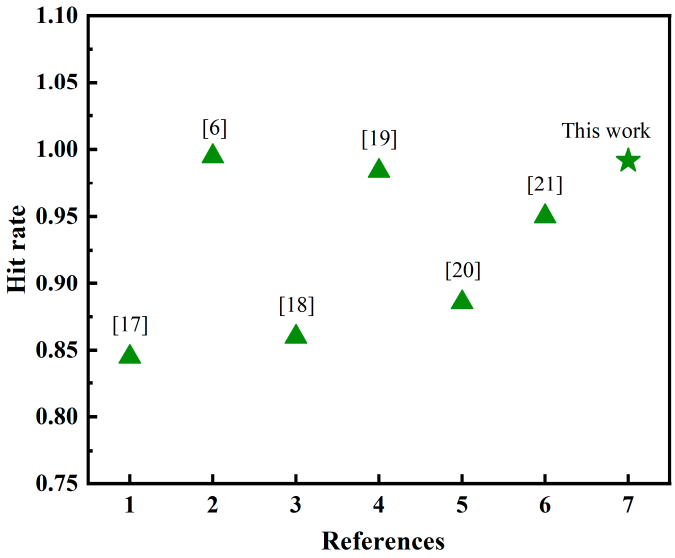

Figure 10. Comparison of prediction accuracy of alloy yields among models [6,17–21].

In summary, the neural network model is more effective than the traditional simple prediction model in predicting non-linear data and is suitable for fitting the elemental yield of the steelmaking process. A comparison between the BP neural network and the LSTM neural network shows that the latter provides more accurate predictions. The WOA-LSTM model and the t-SNE-WOA-LSTM combined model can improve the fitting ability of the prediction model based on the LSTM neural network.

3.4. Application Effect Evaluation

The t-SNE-WOA-LSTM Mn/Si element yield prediction model established in this study was applied to the Q345B production process. Production data were collected to verify the application effect of the prediction model and to compare the element content of the steel after adding ferroalloys according to the predicted yield. A sample of production data from 20 furnaces was used for comparative analysis. The statistical results of actual application errors are shown in Figure 11a. The yield prediction error for Mn in the 20 test furnaces is within ±3%. Figure 11b displays a comparison of the Mn/Si element content in the final liquid steel product. The results demonstrate that the element content control of the finished product is more stable and accurate than the traditional method.

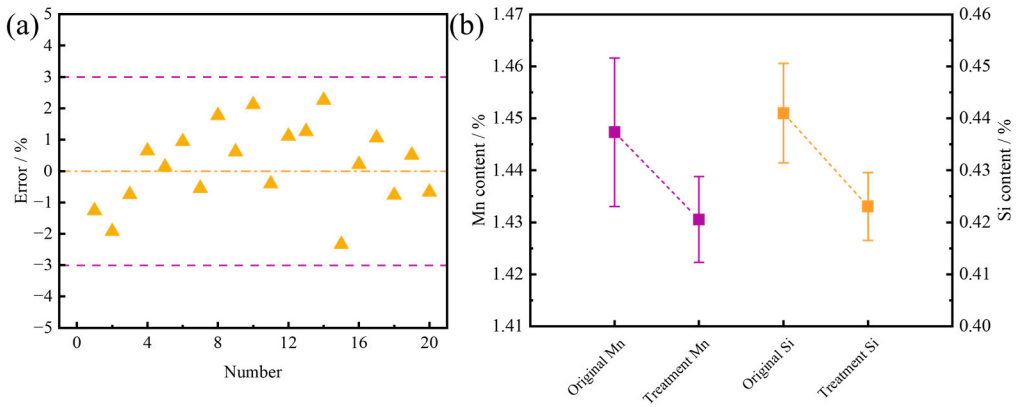

Figure 11. (**a**) Application error; (**b**) Mn/Si content in molten steel.

4. Conclusions

This paper focuses on the Mn/Si element yield in the Q345B smelting process. The element yield is predicted using an LSTM neural network combined with correlation analy-

sis, the t-SNE dimensionality reduction algorithm, and the WOA optimization algorithm. The research and analysis lead to the following conclusions:

(1) The steelmaking process element yield is time series data, which is related to the current furnace influencing factors, and the data are non-linear in character. It is appropriate to use the LSTM neural network to predict the Mn/Si elemental yield. The Mn content and the amount of silicomanganese alloy show a strong negative correlation; the lower the Mn content, the greater the amount of the silicomanganese alloy.

(2) The optimal t-SNE-WOA-LSTM neural network structure was obtained by comparing the prediction results under different experimental hyperparameters. The prediction model for Mn yield has a learning rate of 0.01, a dropout rate of 0.05, a batch size of 32, 194 iterations, 30 neurons, and a time step of 4. The prediction model for Si yield has a learning rate of 0.0043, a dropout rate of 0.4621, a batch size of 16, 89 iterations, 56 neurons, and a time step of 3.

(3) This study compared the t-SNE-WOA-LSTM neural network structure with other prediction models. The Mn yield model's prediction error ranges were within ±1%, ±2%, and ±3%, respectively, with a prediction accuracy of 71.67%, 96.67%, and 99.17%. The Si yield model's prediction error ranges were within ±1%, ±3%, and ±5%, respectively, with a prediction accuracy of 44.17%, 75.83%, and 91.67%. The t-SNE-WOA-LSTM neural network structure had the best prediction accuracy among the models.

(4) During practical application, the element yield model has a prediction error within ±3%. The Mn/Si content in the liquid steel is at the lower limit of the control standard, achieving 'narrow composition' control of the liquid steel.

Author Contributions: Methodology, Y.B.; Software, L.Z.; Validation, X.L., Y.B. and L.Z.; Investigation, X.L.; Resources, X.Q., X.X. and S.L.; Data curation, X.L., X.Q., X.X. and S.L.; Writing—original draft, X.L.; Writing—review & editing, X.L.; Project administration, Y.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (52174297). The authors wish to express their gratitude to the foundation for providing financial support.

Data Availability Statement: The data used in this article is classified.

Conflicts of Interest: Authors Xihui Qu, Xinjun Xie and Sijun Li were employed by the company Shan Dong Iron and Steel Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Fan, Z.Y.; Friedmann, S.J. Low-carbon production of iron and steel: Technology options, economic assessment, and policy. *Joule* **2021**, *5*, 829–862. [CrossRef]
2. Zhang, X.Y.; Jiao, K.X.; Zhang, J.L.; Guo, Z.Y. A review on low carbon emissions projects of steel industry in the World. *J. Clean. Prod.* **2021**, *306*, 127259. [CrossRef]
3. Zhang, Y.Y.; Yu, L.H.; Cui, K.K.; Wang, H.; Fu, T. Carbon capture and storage technology by steel-making slags: Recent progress and future challenges. *Chem. Eng. J.* **2023**, *455*, 140552. [CrossRef]
4. Ren, L.; Zhou, S.; Peng, T.D.; Ou, X.M. A review of CO_2 emissions reduction technologies and low-carbon development in the iron and steel industry focusing on China. *Renew. Sustain. Energy Rev.* **2021**, *143*, 110846. [CrossRef]
5. Zheng, R.X.; Bao, Y.P.; Wang, Z.L. Intelligent control model of steelmaking using ferroalloy reduction and its application. *Chin. J. Eng.* **2021**, *43*, 1689–1697.
6. Xin, Z.C.; Zhang, J.S.; Jin, Y.; Zheng, J.; Liu, Q. Predicting the alloying element yield in a ladle furnace using principal component analysis and deep neural network. *Int. J. Miner. Metall. Mater.* **2023**, *30*, 335–344. [CrossRef]
7. Zeng, J.Q. Thoughts on intellectualization improvement of iron and steel production process. *Met. Ind. Autom.* **2019**, *43*, 1.
8. Zhou, C.G.; Hu, J.Z.; Jiang, C.M.; Wang, S.H.; Ai, L.Q.; Chen, H. Prediction model of phosphorus content in dephosphorization converter end point based on BP neural network algorithm. *Steelmaking* **2021**, *37*, 10–15.
9. Wu, S.W.; Yang, J. A convolutional neural network-based model for predicting lime utilization ratio in the KR desulfurization process. *Metall. Res. Technol.* **2021**, *118*, 603. [CrossRef]
10. Xie, Q.; Suvarna, M.; Li, J.L.; Zhu, X.Z.; Cai, J.J.; Wang, X.N. Online prediction of mechanical properties of hot rolled steel plate using machine learning. *Mater. Des.* **2021**, *197*, 109201. [CrossRef]

11. Acosta, S.M.; Amoroso, A.L.; Sant'Anna, Â.M.O.; Junior, O.C. Predictive modeling in a steelmaking process using optimized relevance vector regression and support vector regression. *Ann. Oper. Res.* **2022**, *316*, 905–926. [CrossRef]
12. Xie, T.Y.; Zhang, C.D.; Zhou, Q.L.; Tian, Z.Q.; Liu, S.; Guo, H.J. TSC prediction and dynamic control of BOF steelmaking with state-of-the-art machine learning and deep learning methods. *J. Iron Steel Res. Int.* **2024**, *31*, 174–194. [CrossRef]
13. Zheng, R.X.; Bao, Y.P.; Zhao, L.H.; Xing, L.D. Method to predict alloy yield based on multiple raw material conditions and a PSO-LSTM network. *J. Mater. Res. Technol.* **2023**, *27*, 3310–3322. [CrossRef]
14. Marani, M.; Zeinali, M.; Songmene, V.; Mechefske, C.K. Tool wear prediction in high-speed turning of a steel alloy using long short-term memory modelling. *Measurement* **2021**, *177*, 109329. [CrossRef]
15. Xu, Z.; Mao, Z.Z. Analysis and prediction of influencing factor on element recovery in ladle furnace. *Iron Steel* **2012**, *47*, 34–37. [CrossRef]
16. Song, C.N.; Cao, J.G.; Wang, L.L.; Xiao, J.; Zhao, Q.F. The prediction model for transverse thickness difference of electric steel in 6-high cold rolling mills based on GA-PSO-SVR approach. *Steel Res. Int.* **2022**, *93*, 2200302. [CrossRef]
17. Nath, N.K.; Mandal, K.; Singh, A.K.; Basu, B.; Bhanu, C.; Kumar, S.; Ghosh, A. Ladle furnace on-line reckoner for prediction and control of steel temperature and composition. *Ironmak Steelmak* **2006**, *33*, 140–150. [CrossRef]
18. Yi, Z.; Chai, L.; Liu, H.K.; Yang, L. Prediction model of C and Mn alloy yield in LF based on AO-ENN. *China Metall.* **2022**, *32*, 40–48+57.
19. He, X.Y.; Wang, M.; Ji, J.L.; Bao, Y.P.; Yang, J.F.; Wang, Z.L. Prediction of Mn alloying yield in converter tapping process based on GA-BP neural network. *Steelmaking* **2022**, *38*, 14–20.
20. Han, M.; Xu, Q.; Zhao, Y.; Lin, D.; Yang, X.L. Calculation of alloy addition to yied-predict model BOF steelmaking. *Steelmaking* **2010**, *26*, 44–47.
21. Zhou, K.X.; Lin, W.H.; Sun, J.K.; Feng, X.M.; Fang, W.; Liu, Q. Aprediction model to calculate Mn yield during BOF alloying process using improved extreme learning machine. *J. Cent. South Univ. (Sci. Technol.)* **2021**, *52*, 1399–1406.
22. Liu, X.; Bao, Y.P.; Zhao, L.H.; Gu, C. Establishment and Application of Steel Composition Prediction Model Based on t-Distributed Stochastic Neighbor Embedding (t-SNE) Dimensionality Reduction Algorithm. *J. Sust. Metall.* **2024**. [CrossRef]
23. Kobak, D.; Linderman, G.C. Initialization is critical for preserving global data structure in both t-SNE and UMAP. *Nat. Biotechnol.* **2021**, *39*, 156–157. [CrossRef] [PubMed]
24. Kang, B.; Garcia, D.; Lijffijt, J.; Santos-Rodríguez, R.; De, T. Conditional t-SNE: More informative t-SNE embeddings. *Mach. Learn.* **2021**, *110*, 2905–2940. [CrossRef] [PubMed]
25. Liu, Y.J.; Dong, N.; Wang, Z.Y. A Fault-Tolerant Soft Sensor Algorithm Based on Long Short-Term Memory Network for Uneven Batch Process. *Processes* **2024**, *12*, 495. [CrossRef]
26. Song, X.M.; Yin, F.B.; Zhao, D.F. Fused Data-Driven Approach for Early Warning Method of Abnormal Conditions in Chemical Process. *Processes* **2023**, *11*, 2435. [CrossRef]
27. Ouladbrahim, A.; Belaidi, I.; Khatir, S.; Magagnini, E.; Capozucca, R.; Wahab, M.A. Experimental crack identification of API X70 steel pipeline using improved Artificial Neural Networks based on Whale Optimization Algorithm. *Mech. Mater.* **2022**, *166*, 104200. [CrossRef]
28. Eze, E.; Ajmal, T. Dissolved oxygen forecasting in aquaculture: A hybrid model approach. *Appl. Sci.* **2020**, *10*, 7079. [CrossRef]
29. Tanvir, M.H.; Hussain, A.; Rahman, M.T.; Ishraq, S.; Zishan, K.; Rahul, S.T.T.; Habib, M.A. Multi-objective optimization of turning operation of stainless steel using a hybrid whale optimization algorithm. *J. Manuf. Mater. Process.* **2020**, *4*, 64. [CrossRef]
30. Afzal, S.; Ziapour, B.M.; Shokri, A.; Shakibi, H.; Sobhani, B. Building energy consumption prediction using multilayer perceptron neural network-assisted models; comparison of different optimization algorithms. *Energy* **2023**, *282*, 128446. [CrossRef]
31. Shi, C.Y.; Wang, B.S.; Guo, S.Y.; Sun, P.; Ma, Z.C.; Wang, Y.K.; Zhang, L.; Wu, C.L. Energy Consumption Prediction of Steelmaking Process Based on Improved Whale Optimization Algorithm and Stochastic Configuration Network. *JOM* **2023**, *75*, 4320–4331. [CrossRef]