Research Article

Zhao Liu, Shusen Cheng*, and Pengbo Liu

Prediction model of BOF end-point P and O contents based on PCA–GA–BP neural network

Abstract: Low-carbon, green and intelligent production is urgently needed in China’s iron and steel industry. Accurate prediction of liquid steel composition at the end of basic oxygen furnace (BOF) plays an important role in promoting high-quality, high-efficiency and stable production in steelmaking process. A prediction model based on the principal component analysis (PCA) – genetic algorithm (GA) – back propagation (BP) neural network is proposed for BOF end-point P and O contents of liquid steel. PCA is used to eliminate the correlation between the factors, and the obtained principal components are seen as input parameters of the BP neural network; then, GA is employed to optimize the initialized weights and thresholds of the BP neural network. The flux composition and bottom blowing are considered in the input variables. The results indicate that the prediction accuracy of the single output model is higher than that of the dual output model. The root-mean-square error of P content between predicted and actual values is 0.0015%, and that of O content is 0.0049%. Therefore, the model can provide a good reference for BOF end-point control.

Keywords: BOF, prediction model, principal component analysis, BP neural network, genetic algorithm

1 Introduction

Basic oxygen furnace (BOF) process is the most important steelmaking method in China. Accurately predicting the liquid steel composition at the end-point of the BOF is of great importance to achieve emission reduction, consumption reduction, green and efficient production. The phosphorus and oxygen contents of the liquid steel are a very important factor for end-point control of BOF, which has a significant impact on product quality and plant economics. Since the process after BOF has no dephosphorization function, phosphorus must be removed to the desired level in the BOF process. Phosphorus concentrations below 0.01% may generally be required to avoid embrittlement [1]. Many steel plants have adopted high-efficiency dephosphorization technology in BOF due to the increasing demand for low and ultra-low phosphorus steels [2–4]. In addition, the oxygen content of the liquid steel at the end-point has a significant influence on the alloy consumption, desulfurization and inclusion removal of the subsequent processes. In summary, if the content of phosphorus and oxygen at the end-point of BOF can be controlled more accurately, the quality of liquid steel will be further improved.

The physical and chemical reactions with a high temperature in BOF and the fluctuation of raw materials composition make it still a challenge to accurately predict the composition of the liquid steel at the end-point. At the same time, there exists a complex non-linear relationship between the variables affecting the liquid steel composition at the end of BOF. Therefore, linear regression models and thermodynamic-based mathematical models are difficult to obtain better prediction results. Artificial neural network (ANN) has the capability of self-study, self-adaptation, and self-organization, which can simulate complicated functions. At present, some prediction methods based on neural networks or neural network combined algorithms have been widely used in BOF end-point control [5]. For example, Wang et al. [6] applied weighted K-means clustering and a group method of data handling (GMDH), combined with a separate neural network for each cluster. Using this method, they reported a phosphorus prediction accuracy of 88.40%, with an error tolerance of ±0.004%. He and Zhang [7] proposed a model based on principal component analysis (PCA) and back-propagation (BP) neural network for predicting the BOF end-point phosphorus content. They achieved hit rates
of 92.06 and 82.54% within error bond $\pm 0.004$ and $\pm 0.003\%$, which is higher than those of multiple linear regression and ordinary BP neural networks. Bae et al. [8] examined the prediction accuracy of seven standard machine learning algorithms using a large production data set with the aim of predicting the BOF end-point. The results showed that ANN has better predictive performance than other algorithms and gets an 89% hit rate with an error of $\pm 0.003\%$. Zhou et al. [9] established a prediction model of end-point phosphorus content for BOF based on a monotone-constrained BP neural network. The model using the constraint condition achieved a prediction accuracy of 74% with an error of $\pm 0.003\%$. Zhou et al. [10] selected seven input variables according to the grey correlation analysis as the input of the BP neural network. The authors claimed the hit rate to be 83.33% when the range of phosphorus error is $\pm 0.004\%$. It can be seen that the ANN with high prediction accuracy is widely used for end-point phosphorus content prediction. Moreover, there are many prediction models of liquid steel temperature, carbon content and manganese content based on intelligent algorithms in previous studies [11–16].

However, few works are investigating the prediction of the oxygen content of liquid steel. In addition, researchers usually considered the weight of the flux in the input variables and ignored the chemical composition of the flux. To enhance the accuracy of predicting the BOF end-point phosphorus and oxygen contents of liquid steel, a hybrid model based on PCA, genetic algorithm (GA), and BP neural network is proposed. PCA is used to reduce the dimensionality of the input variables and eliminate the collinearity among the variables, rather than deleting variables with a high coincidence degree [17]. GA is employed to optimize the initialized weights and thresholds of the BP neural network to overcome BP’s disadvantage of being easily stuck in a local minimum. Therefore, the PCA–GA–BP neural network has the ability to search globally and approach any nonlinear functions with any precision.

2 Model establishment

2.1 Input and output parameters

The schematic diagram of BOF steelmaking is shown in Figure 1. During the process, oxygen is used to react with the elements in the liquid steel, and the components in the flux react with the oxidation products to obtain suitable slag, which contains necessary basicity, MgO content and FeO content. At the same time, the bottom blowing process can enhance the stirring of the metal pool and reduce the equilibrium oxygen content in liquid steel.

According to the metallurgical mechanism and production practice, end-point liquid steel composition is mainly determined by the factors listed in Table 1. It should be noted that the fluctuation of flux composition will affect the stability of the steelmaking process, so the chemical composition of the flux must be considered. Therefore, the chemical composition carried in the flux is considered part of the input variables. For example, CaO addition is the sum of the CaO in lime, limestone, dolomite and light-burned dolomite. CO$_2$ is the product of carbonate decomposition in flux. Additionally, the bottom blowing process, which can significantly reduce the dissolved oxygen content of liquid steel [18], is set as input to the model. The output variables of the model are the P and O contents of liquid steel.

2.2 PCA

PCA is probably the most popular multivariate statistical technique, and it is used by almost all scientific disciplines. Its goal is to extract the important information from the table, represent it as a set of new orthogonal variables called principal components and display the pattern of similarity of the observations and the variables as points in maps [19]. Principal components are obtained as linear combinations of the original variables, and the process has the following five steps:

1) List and standardize a matrix of observed sample variables. The original variable matrix and the normalized matrix are $X$ and $x$, respectively, as shown in equations (1) and (2).
Table 1: Input and output variables of the model

| Variables                      | Symbols | Units   | Minimum   | Maximum   | Mean   | Standard deviation |
|-------------------------------|---------|---------|-----------|-----------|--------|--------------------|
| Charged hot metal             | $X_1$   | $t$     | 223.30    | 272.50    | 249.68 | 4.19               |
| Hot metal temperature         | $X_2$   | °C      | 1.159     | 1.430     | 1340.32| 41.67              |
| $[C]$ in hot metal            | $X_3$   | %       | 4.10      | 4.65      | 4.45   | 0.08               |
| $[Si]$ in hot metal           | $X_4$   | %       | 0.08      | 1.16      | 0.34   | 0.11               |
| $[Mn]$ in hot metal           | $X_5$   | %       | 0.03      | 0.19      | 0.13   | 0.02               |
| $[P]$ in hot metal            | $X_6$   | %       | 0.05      | 0.15      | 0.12   | 0.01               |
| $[S]$ in hot metal            | $X_7$   | %       | 0.001     | 0.099     | 0.020  | 0.020              |
| Charged scrap                 | $X_8$   | $t$     | 12.84     | 40.90     | 24.79  | 5.34               |
| $O_2$ consumption             | $X_9$   | Nm$^3$  | 9.289     | 13,988    | 11,257 | 549.45             |
| CaO addition                  | $X_{10}$ | $t$     | 7.18      | 18.52     | 10.53  | 1.73               |
| MgO addition                  | $X_{11}$ | $t$     | 0.98      | 4.33      | 1.82   | 0.33               |
| SiO$_2$ addition              | $X_{12}$ | $t$     | 0.16      | 0.44      | 0.24   | 0.04               |
| CO$_2$ addition               | $X_{13}$ | $t$     | 0.75      | 5.19      | 2.10   | 0.73               |
| CaF$_2$ addition              | $X_{14}$ | $t$     | 0.00      | 0.15      | 0.01   | 0.04               |
| $N_2$ consumption             | $X_{15}$ | Nm$^3$  | 2.77      | 475.83    | 55.21  | 26.71              |
| Ar consumption                | $X_{16}$ | Nm$^3$  | 20.56     | 340.66    | 133.49 | 33.05              |
| End-point P content           | $Y_1$   | %       | 0.004     | 0.084     | 0.013  | 0.014              |
| End-point O content           | $Y_2$   | %       | 0.026     | 0.092     | 0.047  | 0.008              |

$X = [X_1, X_2 \ldots X_p] = \begin{bmatrix} X_1 & X_2 & \ldots & X_p \\ X_2 & X_2 & \ldots & X_p \\ \vdots & \vdots & \ddots & \vdots \\ X_n & X_n & \ldots & X_p \end{bmatrix}$, \hspace{1cm} (1)

$x = [x_1, x_2 \ldots x_p] = \begin{bmatrix} x_1 & x_2 & \ldots & x_p \\ x_2 & x_2 & \ldots & x_p \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_n & \ldots & x_p \end{bmatrix}$, \hspace{1cm} (2)

where $x_i = \frac{x_i - \bar{x}_i}{S_i}$, $\bar{x}_i = \frac{1}{n}\sum_{i=1}^{n}X_{ij}$, and $S_i = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \bar{x}_i)^2}$.

The mean of each column of the normalized matrix $x$ is 0, and the variance is 1.

2) Calculate the covariance matrix $R$.

$R = \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1p} \\ r_{21} & r_{22} & \ldots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \ldots & r_{np} \end{bmatrix}$, \hspace{1cm} (3)

where $r_{ij} = \frac{1}{n-1}\sum_{k=1}^{n}(x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)$, $\bar{x}_i = \frac{1}{n}\sum_{i=1}^{n}x_{ij}$, and $\bar{x}_j = \frac{1}{n}\sum_{j=1}^{n}x_{ij}$.

The resulting covariance matrix is the correlation coefficient matrix of the original variables.

3) Compute the eigenvalues $\lambda_i$ of the matrix $R$ and the corresponding orthogonalization unit feature vector $p_i$.

Then arrange eigenvectors in order of size, namely, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$.

4) Calculate the contribution rate and cumulative contribution rate of the principal components, which can be calculated using equations (4) and (5), respectively.

$$c_i = \frac{\lambda_i}{\sum_{k=1}^{p}\lambda_k}, \hspace{1cm} (4)$$

$$C = \frac{\sum_{k=1}^{i}\lambda_k}{\sum_{k=1}^{p}\lambda_k}. \hspace{1cm} (5)$$

5) Calculate principal components. The $i$th principal component of the original sample is as follows:

$$F_i = p_i x_i. \hspace{1cm} (6)$$

2.3 BP neural network model

BP algorithm is the most widely used ANN algorithm currently [20], which is simply a gradient descent method designed to minimize the total error (or mean error) of the output computed by the network. Figure 2 shows the structure of the BP neural network. In this network, there is an input layer, an output layer, and one or more hidden layers.

As the values of the input and output variables may be in several orders of magnitude, the difference would greatly affect the convergence speed and accuracy of the BP neural network. Therefore, the data need to be normalized, and the method is as follows:

$$x'_{ij} = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})}, \hspace{1cm} (7)$$

where $x_{ij}$ and $x'_{ij}$ are the normalized value before and after; $\min(x_{ij})$ and $\max(x_{ij})$ are the minimum and
maximum values of $x_{ij}$, respectively. It is important to note that the training and test sets are normalized using the same rules. The network training is a process of continual readjustment between the weights and the threshold, in order to make the network error reduce to a pre-set minimum or stop at a pre-set training step, input the forecasting samples to the trained network, and then obtain the forecasting results [21].

### 2.4 GA–BP neural network

GA is one of the evolutionary algorithms, which is characterized by population-based evolution, survival of the fittest, directed randomness, and does not rely on gradient information. The traditional network weight training generally uses the gradient descent method; these algorithms get stuck to the local optimum easily and cannot get the global optimum. One aspect of using GA in neural networks is to use GA to learn the connection weights; in this way, we can replace some traditional learning algorithms to overcome their defects [21]. First, use GA to optimize the initial weight distribution and locate some better search spaces in the solution space. Then, a BP algorithm is used to search for the optimal solution in these small solution spaces. Generally, the efficiency of the hybrid training is superior to the training methods that only use a BP evolution or a BP training [22].

A typical algorithm for optimizing the connection weights of the BP neural network using GA is as follows, and the flowchart is shown in Figure 3.

1) Determine the network topology, and initialize weights and thresholds.
2) Initialize a population, including its scale, selection probability, crossover probability and mutation probability. Generate the individual randomly. And each individual of the population represents a group of initial weights of the network.
3) Calculate the fitness of each individual. The fitness calculation method selected in this study is as follows:
   \[ F = \sum_{i=1}^{n} \text{abs}(y_i - o_i), \]  
   where $y_i$ is the predicted value of the $i$th node and $o_i$ is the actual value of the $i$th node.
4) Perform selection, crossover and mutation to find the optimal fitness value corresponding to the individual and generate a new population.
5) Return to step 3, circulate and repeat the above procedure until halting criteria are satisfied.
6) Replace initial weights and thresholds of the BP neural network with the optimized weights computed by the GA, and use the BP algorithm for local search.

### 3 Results and discussion

#### 3.1 Collection and preprocessing of data

In this study, the industrial data of the steelmaking process in 250t BOF, 1293 heats, are collected from a steel plant in China. Because the abnormal data or noise data would disturb the model construction and result in incorrect results [23,24], it is necessary to filter and treat the data to reduce the impact of abnormal data on the model. The $3\sigma$ criterion is used to eliminate abnormal data from the original data [17]. Values with a deviation greater than $3\sigma$ are considered abnormal data. After this criterion, these abnormal data will be removed and 1,184 heats of qualified data were obtained. Ultimately, 1,084 heats are selected to establish the model, 100 heats for model validation. The
training and prediction programs are written with Matlab R2021b based on BP neural network.

### 3.2 Establishment of PCA–BP neural network model

The Pearson correlation coefficients of these variables are calculated using equation (3). Because the covariance matrix of the standardized matrix and the correlation coefficient matrix of the original variables are the same. As shown in Figure 4, there is a collinearity problem among the variables, and the values greater than 0.5 are circled. For the collinearity problem, the main effect is to make the regression coefficients unreliable and increase the variance of regression prediction [25]. PCA was used to remove the effect of correlation between variables on the model. Then, the obtained principal components are used as input variables of the BP neural network.

Figure 5 presents the results of the PCA. The cumulative contribution rate of the ninth principal component reaches more than 85%, so thenine principal components can fully represent the characteristics of the original variables. Therefore, PCA achieves the effect of expressing high dimensional data with low dimensional data.

The correlation coefficients between the principal components are all zero, so PCA achieves the purpose of eliminating the correlation between original variables. The nine principal component equations are as follows:

\[
F_1 = 0.104x_1 + 0.155x_2 + 0.296x_3 + 0.331x_4 + 0.164x_5 + 0.268x_6
- 0.082x_7 - 0.058x_8 - 0.229x_9 + 0.419x_{10} + 0.368x_{11} + 0.414x_{12} \quad (9)
\]

\[
F_2 = -0.200x_1 - 0.277x_2 - 0.065x_3 + 0.016x_4 + 0.118x_5 - 0.042x_6
- 0.375x_7 - 0.391x_8 + 0.398x_9 + 0.053x_{10} - 0.119x_{11} + 0.042x_{12} \quad (10)
\]

\[
F_3 = 0.310x_1 + 0.017x_2 - 0.277x_3 - 0.516x_4 - 0.165x_5
+ 0.212x_6 + 0.131x_7 - 0.070x_8 + 0.246x_9 + 0.251x_{10} + 0.310x_{11} \quad (11)
\]

\[
F_4 = -0.132x_1 + 0.247x_2 - 0.274x_3 + 0.001x_4
+ 0.036x_5 + 0.443x_6 - 0.205x_7 + 0.094x_8 - 0.020x_9 + 0.043x_{10} \quad (12)
\]

\[
F_5 = -0.257x_1 + 0.465x_2 + 0.055x_3 + 0.296x_4 - 0.125x_5 - 0.446x_6
+ 0.094x_7 + 0.443x_8 - 0.205x_9 + 0.094x_{10} - 0.020x_{11} + 0.043x_{12} \quad (13)
\]

\[
F_6 = 0.541x_1 + 0.084x_2 + 0.137x_3 - 0.197x_4 + 0.267x_5 + 0.162x_6
+ 0.049x_7 + 0.294x_8 + 0.184x_9 - 0.093x_{10} - 0.052x_{11} - 0.089x_{12} \quad (14)
\]

\[
F_7 = -0.540x_1 - 0.033x_2 - 0.182x_3 - 0.107x_4 + 0.124x_5 + 0.317x_6
- 0.133x_7 + 0.001x_8 - 0.243x_9 + 0.086x_{10} + 0.260x_{11} + 0.187x_{12} \quad (15)
\]

\[
F_8 = 0.085x_1 + 0.418x_2 - 0.220x_3 + 0.323x_4 + 0.252x_5 + 0.328x_6
- 0.062x_7 - 0.065x_8 - 0.321x_9 + 0.015x_{10} + 0.010x_{11} + 0.074x_{12} \quad (16)
\]

\[
F_9 = -0.229x_1 + 0.151x_2 + 0.381x_3 + 0.396x_4
+ 0.230x_5 + 0.026x_6 - 0.251x_7 + 0.009x_8 - 0.159x_9 - 0.189x_{10}
- 0.812x_1 - 0.210x_2 - 0.070x_3 - 0.076x_{10} + 0.134x_{11} - 0.036x_{12} \quad (17)
\]

It can be determined by PCA that the number of output neurons of the BP neural network is 9. Researchers claim that the networks with a single hidden layer could approximate and continuous function to any desired accuracy [26–28]. Thus, three-layer BP structure model is adopted in the work. In general, the number of neurons in the hidden layer can be determined by the following formula.
to obtain better coordination of network capacity and training time.

\[ n_H = 2n_I + 1, \]

where \( n_H \) is the number of neurons in the hidden layer, and \( n_I \) is the number of neurons in the input layer. To compare the effect of the number of output variables on the results, the BP neural network with PCA of \( 9 \times 19 \times 2 \) and \( 9 \times 19 \times 1 \) were established respectively. After many times of experiments, it is found that the results are better when the activation functions of the hidden layer and output layer are both “tansig.”

### 3.3 Establishment of GA–BP neural network model

From the analysis of Section 2.4, the neural network is first trained by GA and then by BP. The 16 variables in Table 1 are used as the input of the GA–BP neural network. Therefore, the network structure is \( 16 \times 33 \times 2 \) and \( 16 \times 33 \times 1 \). The parameters of the GA are shown in Table 2.

The variation of fitness value with epoch is shown in Figure 6. It can be seen that when the genetic generations of phosphorus and oxygen are 1844 and 1959, respectively, the fitness value no longer changes. In the experiment, the genetic algebra was higher than the above values to achieve higher precision. At the same time, PCA–BP neural network and GA–BP neural network models are also established in this article.

### 3.4 Establishment of PCA–GA–BP neural network model

Through the above discussion, PCA is first used to reduce the dimensionality of the data, then GA is used to optimize the initialized weights and thresholds of the BP neural network, and finally, the structure of the BP neural network with PCA and GA is determined as \( 9 \times 19 \times 2 \) and \( 9 \times 19 \times 1 \).

### 3.5 Prediction results

The models of PCA–BP, GA–BP, and PCA–GA–BP are tested to make a comparison. Table 3 indicates the comparison of phosphorus content prediction accuracy of different models. The comparison between the single output model and the dual output model is also shown in the table. It is obvious that when the error is within \( \pm 0.005\% \), the hit rate of all models is 100%. The prediction accuracy of the model with single output is slightly

| Model                  | Prediction error range (%) |
|------------------------|-----------------------------|
|                        | ±0.005 (%)                  | ±0.004 (%)                  | ±0.003 (%)                  | ±0.002 (%)                  | ±0.001 (%)                  |
| PCA–BP (single)        | 100                         | 96                          | 79                          | 68                          | 43                          |
| PCA–BP (dual)          | 100                         | 96                          | 85                          | 63                          | 34                          |
| GA–BP (single)         | 100                         | 98                          | 85                          | 68                          | 46                          |
| GA–BP (dual)           | 100                         | 94                          | 77                          | 62                          | 36                          |
| PCA–GA–BP (single)     | 100                         | 98                          | 94                          | 84                          | 72                          |
| PCA–GA–BP (dual)       | 100                         | 93                          | 76                          | 62                          | 37                          |
higher than that of the same type of model with a dual output. At the same time, the PCA–GA–BP neural network model with single output exhibits higher prediction accuracy than that of other models.

The prediction accuracy of the oxygen content of different models is shown in Table 4. Similar to the prediction results of phosphorus content, the model with single output has higher accuracy than that with a dual output. Meanwhile, it is obvious that the PCA–GA–BP neural network with single output possesses the highest prediction accuracy.

The reliability is low when the absolute error is used to evaluate the prediction accuracy of the model. For example, when the actual value of phosphorus content is less than 0.004%, the prediction error is 0.002%, and the deviation is more than 50%. Therefore, a relative error can accurately describe the accuracy of the prediction results, which is calculated using the formula:

$$\delta = \frac{\hat{y}_i - y_i}{y_i} \times 100\%,$$

where $y_i$ is the actual value and $\hat{y}_i$ is the predicted value. Only the relative error of the models with single output is compared. The ratios with relative error within ±10% are shown in Figure 7. It can be seen that PCA–GA–BP neural network model has higher prediction accuracy. When the relative error is within ±10%, the hit rates of P and O are 83.0 and 80.0%, respectively.

Moreover, in order to better compare the prediction accuracy, the performances of the models are evaluated by the root-mean-square error (RMSE) defined in equation (20).

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}.$$

The smaller the value of RMSE is, the better the model prediction performance is. The predictive performance of different models is analyzed and compared, as

### Table 4: Comparison of O content prediction accuracy of different models

| Model              | Prediction error range (%) |
|--------------------|-----------------------------|
|                    | ±0.005 (%) | ±0.004 (%) | ±0.003 (%) | ±0.002 (%) | ±0.001 (%) |
| PCA–BP (single)    | 59         | 48         | 37         | 37         | 18         |
| PCA–BP (dual)      | 53         | 46         | 34         | 26         | 15         |
| GA–BP (single)     | 68         | 63         | 54         | 40         | 21         |
| GA–BP (Dual)       | 67         | 61         | 53         | 34         | 16         |
| PCA–GA–BP (single) | 69         | 65         | 49         | 38         | 25         |
| PCA–GA–BP (dual)   | 66         | 59         | 46         | 33         | 12         |

### Table 5: RMSE (10⁻²) comparison of different models

| Model            | Single output | Dual output |
|------------------|---------------|-------------|
|                  | P      | O      | P      | O      |
| PCA–BP           | 0.21   | 0.64   | 0.21   | 0.68   |
| GA–BP            | 0.19   | 0.53   | 0.22   | 0.54   |
| PCA–GA–BP        | 0.15   | 0.49   | 0.22   | 0.58   |

### Table 6: Comparison of prominent methods, deviations, hit rates and RMSE for phosphorus (P)

| Ref.     | Method                        | P (%) | RMSE (%) |
|----------|-------------------------------|-------|----------|
| Wang et al. [6] | Weighted K-means and GMDH | ±0.004 | —        |
| He and Zhang [7] | PCA–BP            | ±0.004 | —        |
| Bae et al. [8]  | ANN              | ±0.003 | —        |
| Zhou et al. [9] | Monotone-constrained BP    | ±0.003 | 0.0030   |
| Zhou et al. [10]| Grey correlation analysis–BP | ±0.004 | —        |
| This study     | PCA–GA–BP        | ±0.003 | 0.0015   |

Figure 7: Distribution of relative error within ±10%.
shown in Table 5. The RMSE of the single output model is lower than that of the dual output model. In other words, when single output is used by the model, its predictive performance is higher than when dual output is used. Therefore, it is advisable to build a single output model for predicting the phosphorus and oxygen contents of liquid steel, respectively.

The comparison between previous research and this study is shown in Table 6. In this study, when the error of phosphorus content is within 0.003%, the hit rate reaches 94%. By comparison, the prediction accuracy of the model in this study is higher than that of previous studies.

4 Conclusions

The present work established a BP neural network model to predict the end-point of BOF steelmaking, which is trained and tested by the data of 250t BOF. The bottom blowing and flux components are considered in the model to make the prediction accuracy higher. The following conclusions can be drawn:

(1) By comparing single output model and dual output model, it is clear that the model with single output shows better predictive performance.

(2) When the relative error is within ±10%, the hit rate of the PCA–GA–BP neural network with a single output for P and O contents reaches 83.0 and 80.0%, respectively.

(3) The simulation experimental results show that the PCA–GA–BP neural network with single output has better regression capability than other models. The RMSE of P content between predicted and actual values is 0.0015%, and that of O content is 0.0049%. Therefore, the model can provide a good reference for BOF end-point control.

Acknowledgments: This work was financially supported by the National Natural Science Foundation of China (Grant no. 62071034).

Funding information: National Natural Science Foundation of China (Grant no. 62071034).

Author contributions: Shusen Cheng: writing the original draft; Zhao Liu: methodology and reviewing the document; Pengbo Liu: resources.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: All authors can confirm that all data used in this article can be published in High Temperature Materials and Processes.

References

[1] Briant, C. L. and S. K. Banerji. Tempered martensite embrittlement in phosphorus doped steels. *Metallurgical Transactions A*, Vol. 10, No. 11, 1979, pp. 1729–1737.

[2] Ogasawara, Y., Y. Miki, Y. Uchida, and N. Kikuchi. Development of high efficiency dephosphorization system in decarburization converter utilizing FeO dynamic control. *ISIJ International*, Vol. 5310, 2013, pp. 1786–1793.

[3] Tomiyama, S., Y. Uchida, H. Mizuno, K. Akiu, and T. Maeda. A novel control algorithm for dephosphorization in an LD converter. *Journal of Process Control*, Vol. 25, 2015, pp. 35–40.

[4] Pal, S. and C. Halder. Optimization of phosphorous in steel produced by basic oxygen steel making process using multi-objective evolutionary and genetic algorithms. *Steel Research International*, Vol. 88, No. 3, 2017, id. 1600193.

[5] Li, L. L., M. Ar, T. Kc, L. Hm, and O. Bj. Rapid prototyping tools for real-time expert systems in the steel industry. *ISIJ International*, Vol. 30, No. 2, 1990, pp. 90–97.

[6] Wang, H. B., A. J. Xu, L. X. Ai, and N. Y. Tian. Prediction of endpoint phosphorus content of molten steel in BOF using weighted K-means and GMDH neural network. *Journal of Iron and Steel Research International*, Vol. 19, No. 1, 2012, pp. 11–16.

[7] He, F. and L. Zhang. Prediction model of end-point phosphorus content in BOF steelmaking process based on PCA and BP neural network. *Journal of Process Control*, Vol. 66, 2018, pp. 51–58.

[8] Bae, J., Y. Li, N. Stähl, G. Mathiason, and N. Kojola. Using machine learning for robust target prediction in a Basic Oxygen Furnace system. *Metallurgical and Materials Transactions B*, Vol. 51, No. 4, 2020, pp. 1632–1645.

[9] Zhou, K. X., W. H. Lin, J. K. Sun, J. S. Zhang, D. Z. Zhang, X. M. Feng, et al. Prediction model of end-point phosphorus content for BOF based on monotone-constrained BP neural network. *Journal of Iron and Steel Research International*, Vol. 29, 2022, pp. 751–760.

[10] Zhou, C. G., J. Z. Hu, C. M. Jiang, S. H. Wang, L. Q. Ai, and H. Chen. Prediction model of phosphorus content in dephosphorization converter end point based on BP neural network algorithm. *Steelmaking*, Vol. 37, No. 2, 2021, pp. 10–15.

[11] Wang, Z., Z. Chang, Q.P. Ju, F. M. Xie, B. Wang, H. W. Li, et al. Prediction model of end-point manganese content for BOF steelmaking process. *ISIJ International*, Vol. 52, No. 9, 2012, pp. 1585–1590.

[12] Li, W., Q. M. Wang, X. S. Wang, and H. Wang. Endpoint prediction of BOF steelmaking based on BP neural network combined with improved PSO. *Chemical Engineering Transactions*, Vol. 51, 2016, pp. 475–480.

[13] Gao, C., M. Shen, X. Liu, L. Wang, and M. Chen. End-point prediction of BOF steelmaking based on KNNWTSVR and
[14] Chen, Z. X., H. Liu, and L. Qi. Feature selection of BOF steelmaking process data by using an improved grey wolf optimizer. *Journal of Iron and Steel Research International*, Vol. 29, No. 8, 2022, pp. 1205–1223.

[15] Wang, Z., Q. Liu, H. Liu, and S. Wei. A review of end-point carbon prediction for BOF steelmaking process. *High Temperature Materials and Processes*, Vol. 39, No. 1, 2020, pp. 653–662.

[16] Gu, M. Q., A. J. Xu, F. Yuan, X. M. He, and Z. F. Cui. An improved CBR model using time-series data for predicting the end-point of a converter. *ISIJ International*, Vol. 61, No. 5, 2021, pp. 2564–2570.

[17] Zhou, H., Y. Hu, B. Wen, S. Wu, M. Kou, and Y. Luo. BP neural network prediction for Si and S contents in hot metal of COREX process based on mathematical analysis and Deng’s correlation. *Metallurgical Research & Technology*, Vol. 118, No. 5, 2009, pp. 589–597.

[18] Kaike, C. Controlling oxygen activity in the molten steel at blowing end-point of BOF steelmaking. *Iron and Steel*, Vol. 44, No. 5, 2009, pp. 27–81.

[19] Abdi, H. and L. J. Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, Vol. 2, No. 4, 2010, pp. 433–459.

[20] Werbos, P. J. *The roots of backpropagation: from ordered derivatives to neural networks and political forecasting*, 1, John Wiley & Sons, Inc., New York, NY, 1994.

[21] Ding, S., C. Su, and J. Yu. An optimizing BP neural network algorithm based on genetic algorithm. *Artificial Intelligence Review*, Vol. 36, No. 2, 2011, pp. 159–162.

[22] Xiangping, M., Z. Huaguang, and T. Wanyu. A hybrid method of GA and BP for short-term economic dispatch of hydrothermal power systems. *Mathematics and Computers in Simulation*, Vol. 51, No. 3–4, 2000, pp. 341–348.

[23] Rivera, W. A. and P. Xanthopoulos. A priori synthetic oversampling methods for increasing classification sensitivity in imbalanced data sets. *Expert Systems with Applications*, Vol. 66, 2016, pp. 124–135.

[24] Chen, J., M. Mahfouf, and G. Sidahmed. A new holistic systems approach to the design of heat treated alloy steels using a biologically inspired multi-objective optimisation algorithm. *Engineering Applications of Artificial Intelligence*, Vol. 37, 2015, pp. 103–114.

[25] Chen, C., N. Wang, and M. Chen. Prediction model of end-point phosphorus content in consteel electric furnace based on PCA-extra tree model. *ISIJ International*, Vol. 61, No. 6, 2021, pp. 1908–1914.

[26] Cybenko, G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, Vol. 2, No. 4, 1989, pp. 303–314.

[27] Hornik, K., M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural networks*, Vol. 2, No. 5, 1989, pp. 359–366.

[28] Hecht-Nielsen, R. *Neurocomputing*, Addison-Wesley, Menlo Park, CA, 1990.