

# Artificial Neural Network for Predicting Silicon Content in the Hot Metal Produced in a Blast Furnace Fueled by Metallurgical Coke

Wandercleiton Cardoso<sup>a\*</sup> , Renzo di Felice<sup>a</sup>, Raphael Colombo Baptista<sup>b</sup>

<sup>a</sup>Università degli Studi di Genova, Department of Civil, Chemical and Environmental Engineering, Genoa, Italy.

<sup>b</sup>Instituto Federal do Espírito Santo, Vitória, ES, Brasil.

Received: August 31, 2021; Revised: November 02, 2021; Accepted: December 03, 2021

The main production route for cast iron and steel is through the blast furnace. The silicon content in cast iron is an important indicator of the thermal condition of a blast furnace. High silicon contents indicate an increase in the furnace's thermal input and, in some cases, may indicate an excess of coke in the reactor. As coke costs predominate in the production of cast iron, tighter control of the silicon content therefore has economic advantages. The main objective of this article was to design an artificial neural network to predict the silicon content in hot metal, varying the number of neurons in the hidden layer by 10, 20, 25, 30, 40, 50, 75, 100, 125, 150, 170 and 200 neurons. In general, all neural networks showed excellent results, with the network with 30 neurons showing the best results among the 12 modeled networks. The validation of the models was confirmed using the Mean Square Error (MSE) and Pearson's correlation coefficient. The cross-validation technique was used to re-evaluate the performance of neural networks. In short, neural networks can be used in practical operations due to the excellent correlations between the real values and those calculated by the neural network.

**Keywords:** Slag, Silicon, Blast furnace, Statistical analysis, Artificial neural network.

## 1. Introduction

The main route for the production of cast iron and steel is through the blast furnace. Steel is generally produced in two steps: First, the pig iron, also called hot metal, is obtained, which consists mainly of iron, a high carbon content (about 4.5%), and impurities such as sulfur, phosphorus, and silicon<sup>1-3</sup>

Hot metal is usually produced in a blast furnace, although a chemically similar but spongy-looking material can be obtained by the process of direct reduction, which is then used in a similar way to the hot metal produced in a blast furnace. However, this article is about the control of silicon in blast furnaces fed with metallurgical coke<sup>4-6</sup>.

The blast furnace is fed from the top with lump iron ore, sinter, pellets, fluxes such as dolomite and limestone, and a fuel called metallurgical coke. Hot air is blown into the furnace from the bottom, through the tuyeres. Fuels such as pulverized coal, biogas and natural gas are also injected into this area. A part of the blast furnace gas is burned in the hot stoves to heat the air jet entering the furnace to about 1050°C. The air jet is enriched with oxygen<sup>7-10</sup>.

The physicochemical interactions gas-solid, solid-solid and liquid-liquid occur in different zones of the blast furnace. The hot air combines with the descending glowing coke to produce carbon monoxide gas and release the energy needed to raise the internal temperature of the blast furnace by about 1650°C<sup>10-14</sup>.

The hot metal and molten slag are removed from the bottom of the furnace and sent to the steel mill for further processing<sup>10</sup>.

However, during hot metal production it is necessary to control the impurities of the hot metal, especially silicon. Hot metal with excess silicon is harmful to the secondary refining process in the steelworks. The final concentration of the hot metal in terms of impurities such as silicon is the result of the equilibrium between the iron phase and the slag phase as the iron droplets percolate through the slag phase<sup>15</sup>.

The silicon content in the cast iron is an important indicator of the thermal condition of a blast furnace and can therefore reflect the quality of the steel. High values of silicon content indicate increased heat input to the furnace and in some cases may indicate excess coke in the furnace. Since coke costs predominate in the production of cast iron, tighter control of silicon content therefore clearly has economic advantages<sup>7,15-19</sup>.

In the field of complex process simulation, the application of solutions based on neural networks has become popular due to their versatility and possibility of development, as well as the greater reliability of the responses, since the neural network receives new data during the training process<sup>20,21</sup>.

Artificial neural networks are computational models that have a number of artificial neurons and mimic the functioning of a human biological neuron. The main feature of this technique is its ability to learn and solve problems that are not linearly separable using information from the environment in which it operates<sup>22-25</sup>.

The basic unit of the neural network is the artificial neuron, which can be divided into input signal, synaptic weight, bias, sum, activation function, and neuron output. These functions mimic the functioning of the human brain

\*e-mail: wandercleiton.cardoso@dicca.unige.it

and make it possible to reproduce the synaptic communication between neurons<sup>2,3,10,26</sup>.

From the analysis of Figure 1, it can be concluded that  $x_j$  represents the input of the network;  $w_{kj}$  represents the synaptic weight, where  $k$  is the neuron number and  $j$  is the input stimulus;  $b_k$  represents the weight parameter (bias);  $f(\cdot)$  represents the activation function of the neuron;  $u_k$  represents the linear combination of the input signals; and  $y_k$  represents the output response of the neuron.

The most common artificial neural network architectures are single-layer feed-forward networks, multilayer feed-forward networks, recurrent networks, and lattice networks<sup>2,27</sup>. In this paper, a single-layer feedforward architecture with a sigmoid activation function and the Levenberg-Marquardt training algorithm was chosen. Figure 2 shows a single-layer feedforward neural network.

In this type of architecture, the input layer is directly connected to one or more neurons that generate the output response. This type of architecture is used to solve classification and pattern problems. Perceptron and Adaline neural networks work with feedforward architecture.

Defining the topology of an artificial neural network (ANN) is not a trivial task, as it is usually determined on the basis of past experience and trial-and-error processes. According to the literature, several authors suggest that one or two layers are sufficient for modeling metallurgical problems. Only one hidden layer, using a sigmoid-type activation function, is sufficient for the network to converge to good results as long as there are enough input variables to train the algorithm<sup>2,16,28</sup>.

The Levenberb-Marquardt (LM) is algorithm for approximating the minimum error function by Newton's method. This approximation is described by Equation 1.

$$\Delta x = [J^T(x)J(x) + \mu I]^{-1} J(x)e(x) \tag{1}$$

Where, ( $I$ ) is the identity matrix, ( $e_{(x)}$ ) represents the error, and ( $J$ ) corresponds to the Jacobian matrix.

Considering all the context presented so far, the main objective of this paper was to model 12 different artificial neural networks to find out what is the ideal number of neurons in the hidden layer that gives the best results for predicting the silicon content in the hot metal. The number

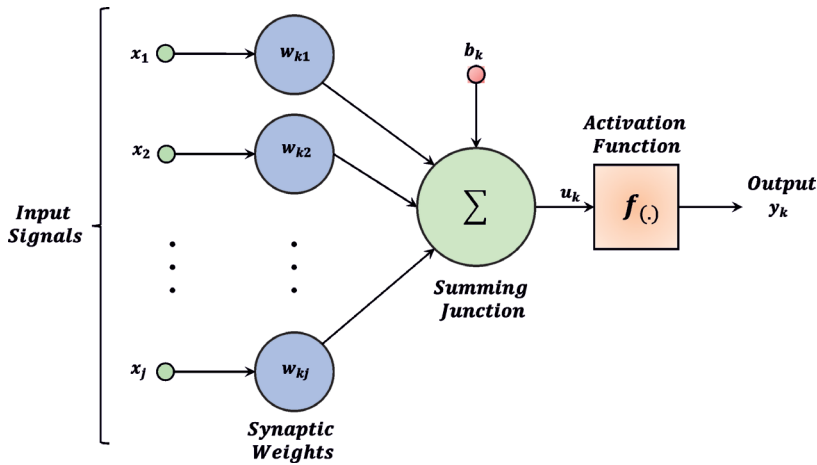


Figure 1. Artificial neuron.

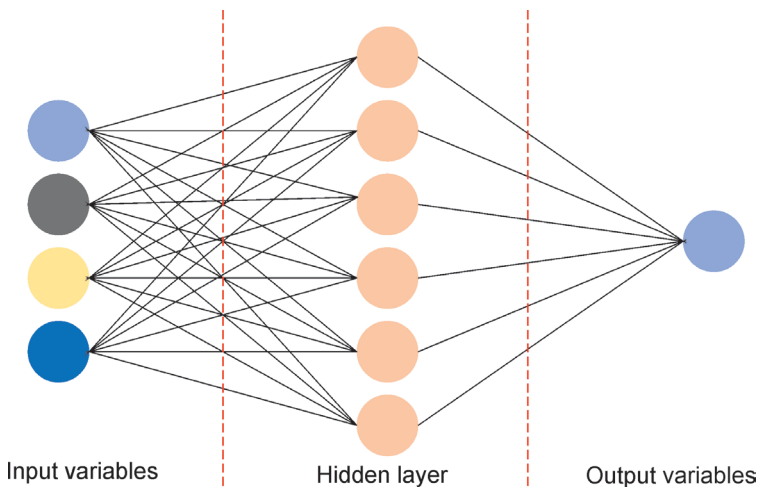


Figure 2. Feed-forward single-layer artificial neural network.

of neurons in the hidden layer were 10, 20, 25, 30, 40, 50, 75, 100, 125, 150, 175, and 200 neurons.

## 2. Research Method

### 2.1. Selection of process variables

The blast furnace studied in this paper has a total height of 110m and a useful height of 43m (considering the useful volume), while its internal volume is 3,617.01m<sup>3</sup>. The crucible of the blast furnace, where the liquid metal and slag are stored, has a volume of 899.03m<sup>3</sup>.

There are a large number of process variables in the blast furnace that affect the silicon content of the hot metal. Selecting an appropriate set of these variables is not a trivial task. The inclusion of many secondary variables in the model complicates its training and use. On the other hand, the inclusion of few variables makes the model poor<sup>2,10,29</sup>.

The input variables were classified into 7 groups. Table 1 shows the summary of the classification of the groups of input variables. Table 2 shows the output variables analyzed.

A total of 1100 operating days over a 3.5-year period were selected. 75 variables were selected as described in Tables 2 and 3. The database consists of 82,500 data.

During these 3.5 years the blast furnace was in normal operation and therefore the output of the neural network corresponds to the normal operation of the blast furnace.

### 2.2. Outlier removal

The database initially consisted of 1302 days of operation and after data processing and removal of the outlier resulted in a database composed of 1100 operational days.

**Table 1.** Variable group division.

Variable groups	Number of variables	Number of information
Blowing air	6	6600
Top gas	7	7700
Temperature	6	6600
Fuel	22	24200
Ore	11	12100
Hot metal	9	9900
Slag	13	14300
<b>TOTAL</b>	<b>74</b>	<b>81400</b>

**Table 2.** Output variable.

Output variable	Number of variables	Number of information
Silicon	1	1100
<b>TOTAL</b>	<b>1</b>	<b>1100</b>

**Table 3.** Operacional outlier.

Operacional outlier	Number of events	Number of days
Blowing air	05	15
Permeability	02	10
Operational maintenance	10	30
Weeks before/after maintenance	20	140
<b>TOTAL</b>	<b>37</b>	<b>195</b>

In this paper, 2 techniques were used to identify the outliers. The first technique used the experience of engineers, consultants and operational technicians.

During a period of 3.5 years, the operation of the blast furnace under study was interrupted 10 times to perform maintenance on the reactor. The maintenance shutdowns lasted on average 3 days (72 hours) each. In the week before and after the maintenance shutdown, operational changes are made at the blast furnace. It was determined that the 7 days before the maintenance shutdown and the 7 days after the maintenance shutdown would be classified as an operational outlier.

During the 3.5 years, the blast furnace experienced 7 serious operational instability events. Five events involved an immediate reduction in the murmur volume and pressure of the reactor. It was determined that the day of the sudden reduction in blow volume event and the 2 days later of event were also classified as operational outlier.

There were also 2 events of load-lowering reduction (permeability) in the blast furnace. It was determined that the day of the event and the 2 days before and after the event are also classified as operational outlier. When the permeability of the blast furnace is reduced, the passage of air through the metallic charge inside the reactor is affected, and when this air is trapped in a zone, the internal pressure increases, resulting in a cage and consequently a decrease in the load of furnace.

Table 3 illustrates the number of days that were considered outliers and excluded from the database.

The second technique to identify the outliers used the principle of exploratory data analysis. The method consists of defining a pair of inner fences and a pair of outer fences, as illustrated in Equations 2 to 5.

The data located between the inner fences is the correct data, those between the inner and outer fences are the moderate outliers and those outside the outer fences are the severe outliers. Where  $Q_1$ =first quartile;  $Q_3$ =third quartile; and  $IQR$ =interquartile range<sup>10</sup>.

$$\text{Lower inner fence} = Q_1 - 1.5 \times (IQR) \quad (2)$$

$$\text{Lower outer fence} = Q_1 - 3 \times (IQR) \quad (3)$$

$$\text{Lower inner fence} = Q_3 - 1.5 \times (IQR) \quad (4)$$

$$\text{Lower outer fence} = Q_3 - 3 \times (IQR) \quad (5)$$

Following the principle of exploratory data analysis, 23 days were found to be moderate outliers and 7 operational days were found to be severe outliers. The 23 days (moderate outliers) were retained in the database, but 7 days considered severe outliers were eliminated.

Table 4 illustrates the composition of the database. Figure 3 shows a normal distribution curve and the interval ( $> +3\sigma$  and  $< -3\sigma$ ) that was considered a severe outlier and removed from the database.

In this paper, considering a normal distribution of the database, all points outside the range ( $\mu + 3\sigma$ ) and ( $\mu - 3\sigma$ ) were considered as outliers and eliminated. In the case of

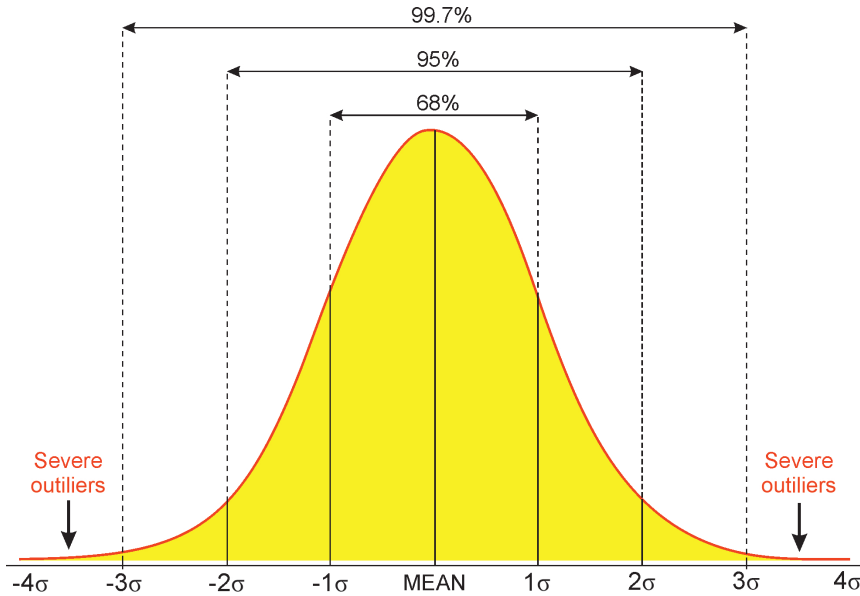


Figure 3. Normal distribution curve.

Table 4. Raw database.

	Number of days	Eliminated	Sustained
<b>Raw database</b>	1302	*	*
Operational outliers	*	-195	*
Moderate outliers	*	*	23
Severe outliers	*	-7	*
<b>Clean database</b>	<b>1100</b> (1302 - 202)	<b>202</b>	<b>23</b>

normally distributed composition data, as in the case of the database of this research, the removal of outliers is justified<sup>10</sup>.

### 2.3. Database standardization

The database contains several physical variables such as temperature, hot metal production, slag production and blown air volume with very different magnitudes.

These data cannot be used directly to train an artificial neural network because variables with a high magnitude affect the respective synaptic weights compared to variables with a lower magnitude. In this paper, all variables were standardized to the range between 0 and 1, as shown in Equation 6.

$$Z = \frac{x - \mu}{\sigma} \quad (6)$$

To standardize a variable ( $Z$ ), the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) are calculated. Thus, for each observed value ( $x$ ) of the variable, the mean is subtracted and divided by the standard deviation. Database standardization is important to interpret the skewness and kurtosis behavior of each variable in the database.

Skewness is the degree of deviation or departure from symmetry of a distribution. When the curve is symmetric, the mean, median, and mode coincide at the same point on the maximum ordinate, so there is perfect equilibrium in

the distribution. If no equilibrium occurs, that is, the mean, median, and mode are at different points in the distribution, it is asymmetric, that is, skewed to the right or left<sup>10</sup>.

A distribution with negative skewness exists when the values are concentrated at the top of the scale and gradually spread to the lower left. The skewness is positive when the third quartile moves away from the median, while the first quartile approaches the median and has a limit:  $Q_1 = Q_2$  when the skewness takes the maximum positive value: ( $S=1$ ).

A distribution with positive skewness exists when the values are concentrated at the top of the scale and gradually spread to the lower right. The skewness becomes negative when the first quartile moves away from the median while the third quartile approaches the median, which is a limit:  $Q_3 = Q_2$  when the skewness takes a maximum negative value<sup>10</sup>.

Kurtosis is the degree of flatness of a distribution, relative to the normal distribution. Kurtosis can be of three types: mesokurtic, when the distribution is normal; leptokurtic, when the distribution is sharper than normal; and platykurtic, when the distribution is flatter than normal<sup>10</sup>. Figure 4 illustrates the behavior of skewness, while Figure 5 illustrates the behavior of kurtosis.

### 2.4. Data segmentation

Several authors argue that normally 85% of the data is used to train and validate the neural network and the remaining 15% of the dataset is presented to the neural network only when its performance in the training and validation phase is considered satisfactory<sup>2,16,30</sup>.

The test step is important to evaluate the generalization and learning ability of the neural network. The neural network must be able to generalize what it has learned and reproduce the solutions from the trained examples for any problems similar to the training<sup>2,16</sup>.

The database of the blast furnace operating records was segmented into 4 groups: (1) training, (2) validation, (3) test, and (4) cross-validation.

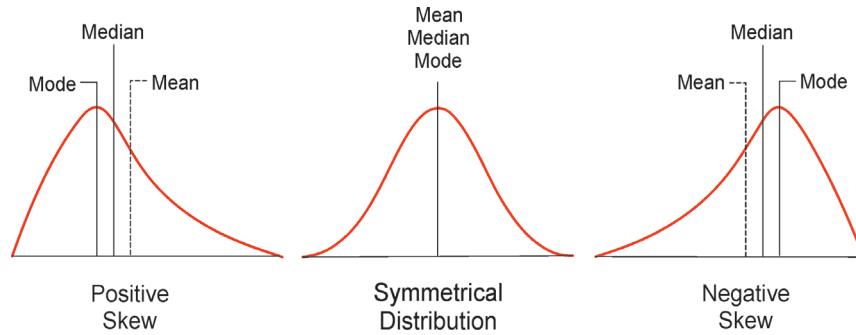


Figure 4. Skewness behavior in a normal distribution.

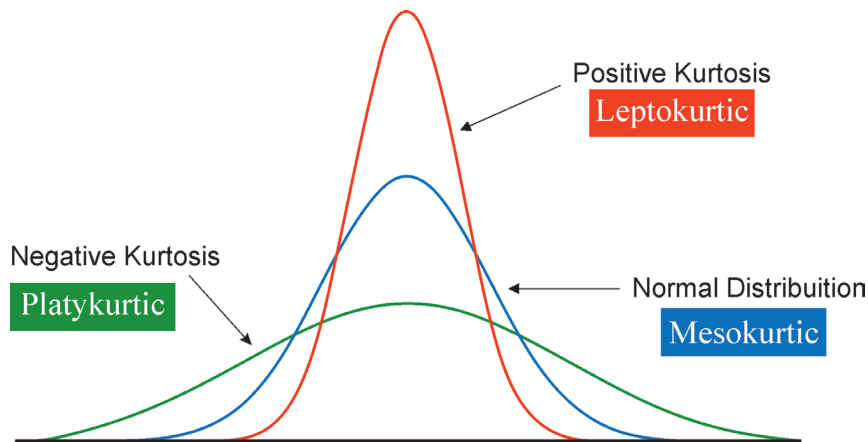


Figure 5. Kurtosis behavior in a normal distribution.

Segmenting the data into training set and validation set ensures that the weights do not converge to a local minimum. In this method, after the neural network has converged with the training data set, it is trained again with the validation data set<sup>10</sup>.

In the validation dataset, the final weights obtained in the training phase are used as the initial weights. If the error converges in the validation phase, the trained network can be used as a process model. To test the accuracy of the neural network model, test data is used. This data is used only once to test the accuracy of the model<sup>31-33</sup>.

After testing, it is possible to check the final result with an additional database to cross-validate the results obtained after the training, validation and testing steps.

The segmentation of the variables was performed randomly and an ANOVA test was performed to verify that the original set and the 4 segmented sets represent the same population given by their mean and standard deviation. Table 5 illustrates the segmentation of the database.

### 2.5. Network architecture

As explained above and shown in Figures 1 and 2, artificial neural networks are known to be computer models inspired by the human brain and used in machine learning and pattern recognition, where the smallest component of a neural network is the artificial neuron.

The optimization algorithm for training was the Levenberg-Marquardt (LM) because it allows for fast convergence.

Table 5. Division of variables.

Step	Variables
Training	700
Validation	150
Test	150
Cross-validation	100
<b>TOTAL</b>	<b>1100</b>

The function used was of the sigmoid type. The model was evaluated using the mean square error (MSE) and the Pearson's correlation coefficient (R).

According to the literature, the number of neurons in the hidden layer must be determined empirically, without explicit rules for an ideal calculation. It is recognized in the literature that in most cases the use of a single hidden layer is sufficient since this structure is able to approximate any non-linear equation (such as quadratic or exponential equations), as long as there are enough input variables to train the neural network.

Two hidden layers are already capable of representing any relationship between the data, including those that cannot be represented by equations.

More than two hidden layers are needed only for even more complex problems, such as time series and computer vision, where there is some relationship between the dimensions contained in the data (time in the first case and geometric shapes in the second).

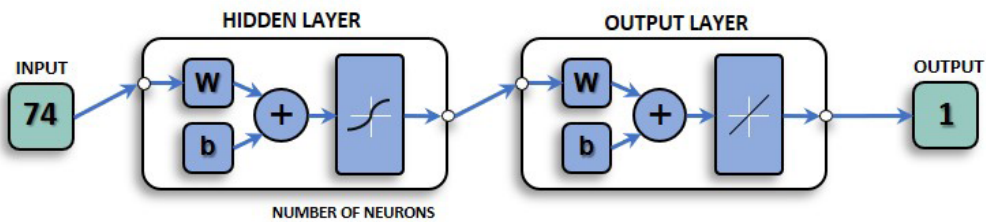
The neural network in this paper has 74 input variables and 1 output variable operating on 82,500 pieces of information. Figure 6 illustrates the architecture of an artificial neural network.

The purpose of this paper is to evaluate the behavior of a neural network with a single hidden layer, sigmoid activation function and using Levenberg-Marquardt algorithm varying the number of neurons in the hidden layer in 10, 20, 25, 30, 40, 50, 75, 100, 125, 150, 175 and 200 neurons.

## 2.6. Statistical analysis

Descriptive statistics applies techniques to describe and summarize a database. Some measures commonly used to summarize a database are measures of central tendency and measures of variability or dispersion.

Measures of central tendency include mean, median, and mode. Measures of variability include standard deviation, maximum and minimum values, skewness, and kurtosis.



**Figure 6.** Artificial neural network architecture.

**Table 6.** Top gas variables.

Variable	Unit	Mean	Std_dev	Minimum	Median	Maximum	Skewness	Kurtosis
CO	%	23.8	0.74	21.1	23.91	26.71	-0.56	0.31
CO <sub>2</sub>	%	24.3	0.66	21.9	24.41	26.10	-0.47	0.21
N <sub>2</sub>	%	47.2	1.39	44.8	46.83	53.87	1.00	1.21
H <sub>2</sub>	%	4.5	0.43	3.0	4.53	5.85	-0.25	0.33
CO + CO <sub>2</sub>	%	47.9	0.6	44.0	48.1	48.1	-3.09	10.68
CO efficiency	%	49.5	0.85	46.07	49.51	52.36	0.00	0.10
H <sub>2</sub> efficiency	%	40.7	3.25	12.80	40.80	51.64	-0.95	5.88

**Table 7.** Blow air variables.

Variable	Unit	Mean	Std_dev	Minimum	Median	Maximum	Skewness	Kurtosis
Volume	Nm <sup>3</sup> /min	4852.90	148.60	4251.20	4854.10	5395.0	-0.21	0.71
Pressure	Kgf/cm <sup>2</sup>	3.87	0.10	3.44	3.89	4.10	-0.49	0.33
Moisture	kg/m <sup>3</sup>	19.81	3.73	12.07	19.80	27.93	0.03	-0.71
O <sub>2</sub> enrichment	%	5.27	0.95	2.91	5.27	6.89	-0.40	-0.64
Steam	%	1.51	1.01	0.10	1.39	6.63	1.22	3.10
Total comsump.	Nm <sup>3</sup> /min	7030.30	213.60	6112.80	7030.30	7824.40	-0.11	0.82
Volume	Nm <sup>3</sup> /min	4852.90	148.60	4251.20	4854.10	5395.0	-0.21	0.71

**Table 8.** Temperature variables.

Variable	Unit	Mean	Std_Dev	Minimum	Median	Maximum	Skewness	Kurtosis
Hot metal	°C	1508.3	12.2	1470.4	1508.7	1542.7	-0.35	0.17
Blowing air	°C	1243.3	13.9	1200.8	1238.3	1249.8	-0.62	-0.86
Top gas	°C	121,35	10,34	90,98	120,98	179,25	0,58	1,50
Flame temperature	°C	2177.6	2108	2071.8	2178.6	2243.2	-0.48	1.03
Slag	°C	1508.3	12.2	1470.4	1508.7	1542.7	-0.35	0.17
Thermal index	-	504.70	54.03	364.77	516.41	659.00	-0.32	-0.52

As mentioned earlier, the segmentation of the database was done randomly and an ANOVA test was performed to verify that the original set and the segmented sets represent the same population given by their mean and standard deviation.

The mean, standard deviation, minimum, median, maximum, asymmetry and kurtosis were calculated. Tables 6 to 12 show the descriptive statistics for the groups of input variables. Table 13 shows the descriptive statistics for the output variable (silicon).

## 3. Result and Discussion

### 3.1. Model variables

In selecting of input variables, it was decided to select the most important variables that affect the operation of the blast furnace. The variables were selected considering 7 groups (Table 1): (1) blowing air; (2) blast furnace gas;

(3) thermal control; (4) fuel; (5) ore, sinter and pellets; (6) hot metal and (7) slag.

Blowing air is supplied at the bottom of the blast furnace. Flow measurements are important for process control of the blast furnace and provide information on operating deviations.

Operational control of the blast furnace gas is important information for the application of thermochemical models

and the frequency of its analysis is important to determine specific carbon consumption and specific air flow 2.

Thermal control is important to ensure the performance of the blast furnace and the quality of the final product. The amount of silicon in the hot metal is directly dependent on the temperature of the hot metal and the quality of the minerals fed from the top of the furnace 20.

**Table 9.** Ore variables.

Variable	Unit	Mean	Std_Dev	Minimum	Median	Maximum	Skewness	Kurtosis
Ore/Coque	-	5.10	0.31	3.82	5.15	6.27	-0.66	0.54
Sinter (type 1)	ton	4536.3	884.2	190.2	4709.3	5983.0	-1.62	3.58
Sinter (type 2)	ton	1697.2	1326.2	50.6	1354.8	5297.0	1.00	0.07
Pellet (type 1)	ton	5132.0	1898.3	13.9	5790.7	7803.6	-1.33	0.71
Pellet (type 2)	ton	4813.7	2183.1	53.7	5696.0	7626.7	-0.96	-0.52
Total metal load	ton	12312	670	10010	12238	13506	-0.72	0.28
Raw material rate	Kg/ton	1578.8	15.1	1528.6	1577.9	1618.5	-0.11	0.10
Ore	%	8.9	4.5	0.6	9.0	18.0	-0.05	-0.88
Sinter	%	39.6	2.8	32.9	40.0	49.0	-0.08	-0.39
Pellet	%	51.5	5.1	42.0	51.4	65.0	0.12	-0.75
Ore day	ton/day	12747	703	10170	12845	14093	-0.84	0.67

**Table 10.** Fuel variables.

Variable	Unit	Mean	Std_dev	Minimum	Median	Maximum	Skewness	Kurtosis
Injection PCI	Kg/ton	58.99	6.16	41.89	57.88	72.29	-0.10	-0.32
Gas rate	Kg/ton	-	-	-	-	-	-	-
Coal/O <sub>2</sub> tax	Kg	755.27	75.57	562.99	745.94	933.38	0.09	-0.63
Coal/air tax	-	170.03	74.12	2.15	191.76	265.25	-1.19	0.06
PCI rate	Kg/ton	175.98	15.61	132.16	176.17	219.74	0.08	0.65
Direct reduction	%	23.38	12.41	3.11	29.58	35.21	-0.94	-1.04
PCI tax		1078.30	540.90	12.20	1297.90	1845.60	-0.98	-0.65
Coke total	Kg/ton	1932.20	911.70	23.20	2348.00	3365.20	-1.28	-0.20
Small coke	Kg/ton	294.63	134.86	3.63	334.93	672.10	-0.65	0.19
Coke (type 1)	Kg/ton	210.70	259.80	1.00	137.50	2133.00	3.47	17.81
Coke (type 2)	Kg/ton	742.00	716.00	22.00	219.00	1769.00	0.55	-1.62
Coke (type 3)	Kg/ton	946.00	956.00	39.00	265.00	2597.00	0.57	-1.71
Coke (type 4)	Kg/ton	1878.00	143.00	1777.00	1878.00	1979.00	0.01	0.03
Coke (type 5)	Kg/ton	1327.50	847.60	15.90	1822.00	2649.60	-0.53	-1.56
Moisture	Kg/ton	6.40	1.41	1.29	6.48	13.74	-0.41	1.47
Coke/load	Kg/ton	11.89	9.56	2.11	2.96	42.44	0.10	-1.80
Small coke total	Kg/ton	4.28	0.03	4.26	4.26	4.32	0.44	-2.02
PCI/load	Kg/load	174.74	14.32	16.73	175.30	212.27	-1.78	15.07
Fuel rate/load	Kg/lod	484.08	18.14	362.91	482.35	599.50	2.53	14.09
Coke total/load	Kg/load	24.52	0.89	22.25	24.24	27.80	1.04	1.20
PCI/day	-	1214.40	44.90	1136.60	1222.10	1299.20	0.00	-0.62
Coke rate	Kg/ton	319.68	25.98	272.82	318.84	417.68	0.87	0.64

**Table 11.** Hot metal variables.

Variable	Unit	Mean	Std_Dev	Minimum	Median	Maximum	Skewness	Kurtosis
Estimated production	ton	7789.5	314.5	7014.4	7856.8	8526.8	-0.4	-0.3
Real production	ton	7787.2	324.5	324.5	7836.1	8563.9	-0.4	-0.2
Carbon	%	4.635	0.169	3.788	4.625	5.173	-0.1	0.8
Chrome	%	0.025	0.002	0.019	0.025	0.042	1.3	5.2
Copper	%	0.007	0.001	0.004	0.007	0.009	-1.2	1.7
Sulfur	%	0.023	0.008	0.008	0.021	0.083	1.5	4.8
Phosphorus	%	0.074	0.006	0.055	0.074	0.097	0.3	0.0
Manganese	%	0.29	0.03	0.22	0.29	0.36	0.0	-0.6
Mn (hot metal/slag)	-	0.13	0.22	0.09	0.16	0.7	1.1	3.1

**Table 12.** Slag variables.

Variable	Unit	Mean	Std_Dev	Minimum	Median	Maximum	Skewness	Kurtosis
Slag rate	Kg/ton	246.99	13.74	206.20	247.95	295.72	-0.09	0.45
B2 basicity	-	1.20	0.04	1.04	1.20	1.30	-0.4	0.70
B4 basicity	-	1.07	0.04	0.91	1.07	1.16	-0.6	0.50
Al <sub>2</sub> O <sub>3</sub>	%	10.71	0.62	6.92	10.69	12.54	-0.7	2.70
CaO	%	43.06	1.55	32.50	43.25	45.42	-3.41	15.77
Sulfur	%	1.15	0.14	0.70	1.14	1.80	0.90	2.00
FeO	%	0.42	0.04	0.32	0.42	0.50	-0.10	-0.35
MgO	%	6.83	0.86	4.44	7.05	8.65	-0.30	-1.10
MnO	%	0.31	0.10	0.06	0.30	0.80	0.70	1.60
SiO <sub>2</sub>	%	36.05	1.36	27.30	36.21	39.60	-2.51	10.42
TiO <sub>2</sub>	%	0.58	0.05	0.43	0.58	0.72	-0.27	-0.01
Production	ton	1980.6	190.8	1494.9	1974.1	2466.3	-0.03	-0.38
Mn (slag/hot metal)	-	0.87	0.22	0.30	0.84	0.91	1.1	3.10

**Table 13.** Output variable.

Variable	Unit	Mean	Std_Dev	Minimum	Median	Maximum	Skewness	Kurtosis
Silicon (%)	%	0.337	0.100	0.107	0.315	0.727	0.95	0.87

The control of the blast furnace fuels, which can be injected through the vents or loaded into the blast furnace from the top, is the main group of variables. The blast furnace studied in this paper was operated for 4 years with 7 different types of coal, according Table 10.

The same reactor has a pulverized coal injection system and can also be operated with natural gas injection. Coal and coke play a dual role in the production of hot metal. As a fuel, it enables the high temperatures (about 1,500° Celsius) required to melt the ore to be reached, and as a reducing agent, it combines with oxygen and reduces the ore at high temperature 16.

Mineral control is important to optimize the final quality of the product and reduce the levels of unwanted impurities such as silicon, phosphorus and sulfur. Four years were spent running the blast furnace and parameterizing the chemical composition of the minerals and fuels. The hot metal control is important because it is necessary to ensure the final quality of the steel<sup>10</sup>.

Production of hot metal with high content of silicon, phosphorus and sulfur increases the final cost of the product and makes secondary refining unprofitable<sup>10</sup>.

Slag is obtained by melting and separating the gangue from the raw materials and fluxes. It consists mainly of thermodynamically stable oxides such as MgO, CaO, Al<sub>2</sub>O<sub>3</sub> and SiO<sub>2</sub>, which constitute up to 95 wt% in the slag. The control of hot-metal/slag is important because silicon should preferably be slagged<sup>2,10</sup>.

### 3.2. Statistical analysis

Analyzing the descriptive statistics of the input variables, one can see that the database generally has little noise, low standard deviation, asymmetry, and near zero kurtosis. In terms of silicon (output variable), the behavior is similar to the input variables.

The hypothesis tests were performed using the database as a reference. The hypothesis tests were performed in Minitab

software. The hypothesis test was performed for the mean and for the standard deviation.

The artificial neural network showed excellent results for all neurons. The hypothesis test proved that all sample groups calculated by the neural network are equal in the database considering 99% confidence interval using Welch method.

Table 14 shows the descriptive statistics of the database (actual values) and the values calculated by the artificial neural network.

### 3.3. Model validation

The usual method for evaluating a neural network mathematical model is using the MSE (mean square error). Small MSE values indicate that the model has better predictive ability. The MSE is given by Equation 7.

$$MSE = \frac{1}{n} \sum_{i=1}^n (C_{neural} - C_{real})^2 \quad (7)$$

In many cases, Pearson's correlation coefficient (R) can be used. However, this parameter evaluates the linear relationship between variables. The Pearson correlation coefficient (R) or linear correlation is given by Equation 8.

$$R = \sqrt{\frac{\sum_{i=1}^n (C_{neural} - C_{real})^2}{\sum_{i=1}^n (C_{real} - C_{neural})^2}} \quad (8)$$

Where, (n) represents the number of observations, (C<sub>neural</sub>) represents the value calculated by the artificial neural network and (C<sub>real</sub>) represents the value measured during the blast furnace operation.

The training, validation and testing phase was performed with up to 1,000 iterations and automatically interrupted when it converged to the smallest error. The model was validated using Pearson's correlation coefficient and mean square error. Five correlation coefficients were calculated: (1) training; (2) validation; (3) test; (4) cross-validation;



**Table 14.** Silicon variables (%).

Variable	Mean	Std_Dev	Minimum	Median	Maximum	Skewness	Kurtosis
<b>Database</b>	0.337	0.100	0.107	0.315	0.727	0.95	0.87
10 neurons	0.337	0.097	0.120	0.315	0.719	0.97	0.84
20 neurons	0.336	0.098	0.136	0.316	0.707	0.92	0.69
25 neurons	0.339	0.099	0.107	0.319	0.731	0.91	0.76
30 neurons	0.333	0.099	0.114	0.314	0.702	0.92	0.74
40 neurons	0.337	0.101	0.063	0.318	0.770	0.88	0.90
50 neurons	0.334	0.097	0.125	0.316	0.708	0.82	0.58
75 neurons	0.338	0.099	0.115	0.318	0.725	0.90	0.81
100 neurons	0.338	0.101	0.100	0.316	0.724	0.86	0.72
125 neurons	0.336	0.100	0.091	0.315	0.690	0.86	0.59
150 neurons	0.326	0.094	0.140	0.311	0.690	0.90	0.81
175 neurons	0.338	0.097	0.107	0.317	0.766	1.01	1.34
200 neurons	0.339	0.104	0.121	0.318	0.713	0.94	0.94

and (5) general correlation. The general correlation was calculated using the training, validation, and test variables.

A total of five mean square error values were calculated: (1) training; (2) validation; (3) testing; (4) cross-validation; and (5) general correlation. The mean square error was calculated with the training, validation, and testing variables. Figure 7 to Figure 10 show the results of model validation.

When analyzing the values for the MSE between training, validation, and testing, as shown in Figures 9 and 10, no differences were found that could indicate overfitting, i.e., when the model has a low error during training and a high error during testing.

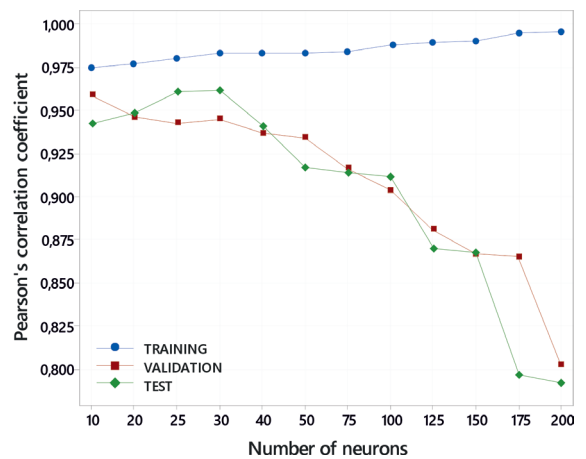
From the analysis of Figures 7, 8, 9 and 10, the best results were obtained with 25 and 30 neurons. The neural network with 30 neurons in the hidden layer showed about 1% better results than the ANN with 25 neurons. From Figures 7 and 8, it can be seen that the neural networks with 25 and 30 neurons have 12% higher mathematical correlation than the network with 200 neurons.

Considering that the best results were those of the neural network with 30 neurons, it can be mentioned that silicon had a Pearson correlation coefficient of 0.975, while the mean square error (MSE) was 0.0006. Regarding cross-validation, it is noted that the best results were also obtained with 25 and 30 neurons. During the cross-validation, the network with 30 neurons showed a mean square error (MSE) of 0.00035 and a Pearson’s correlation coefficient of 0.955.

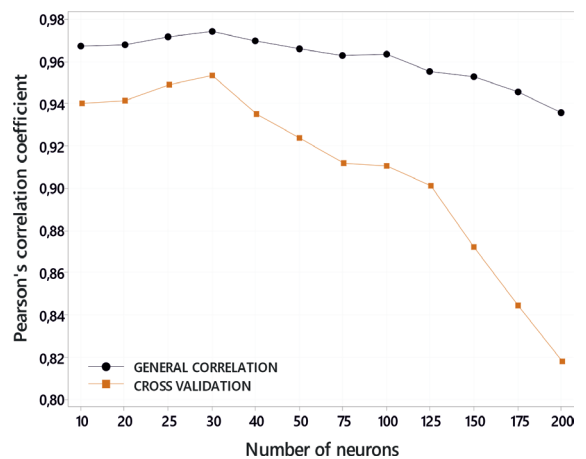
From the analysis of Figure 9, the MSE decreases as the number of neurons increases. In the present study, the neural networks were configured in up to 1000 training epochs. The worst convergence result was obtained by the network with 200 neurons, which required 317 epochs to achieve convergence. The neural network with 30 neurons converged quickly, requiring only 28 epochs to reach convergence.

In this context, the authors Saxén and Pettersson<sup>19</sup> mention that the silicon content has a more irregular behavior, which makes the convergence of the results more difficult. However, this fact was not found in this paper, probably due to the big data used and the elimination of severe outliers, which allowed better learning of the artificial neural network<sup>9,10</sup>.

Based on the results of this research, a comparison was made with the other models mentioned in the literature, as shown in Table 15.



**Figure 7.** Model validation in training, validation and test.

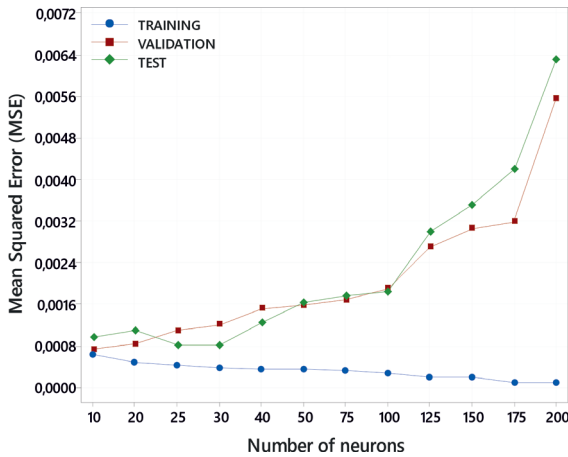
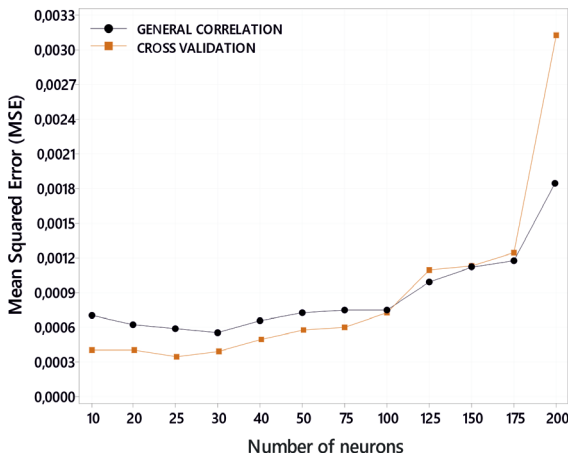


**Figure 8.** Model validation general correlation e cross validation.

The analysis of Table 15 shows that the results of this paper were superior to the models reported in the literature, suggesting that the use of Big Data and the prior treatment of databases is a beneficial alternative in modeling situations to refine the results.

**Table 15.** Comparison between models reported in the literature<sup>16-18</sup>.

	Saxén and Petterson <sup>19</sup>	Dobrzanski et al. <sup>18</sup>	David et al. <sup>16</sup>	Diniz et al. <sup>17</sup>	This paper 30 neurons
MSE	0.0086	0.9695	0.0091	0.0017	0.0006
Pearson (R)	-	-	0.7286	0.9370	0.9752

**Figure 9.** Mean squared error (MSE).**Figure 10.** Mean squared error (MSE).

From a metallurgical point of view, the silicon content in hot metal is an important quality parameter that must be monitored, as this element serves as an indicator of the thermal condition of the reactor. Lower amounts of silicon in the hot metal indicate that the reactor is probably cooling down. On the other hand, an increase in silicon content indicates excessive heat generation and thus wastage of metallurgical coke and pulverized coal<sup>10</sup>.

The silicon in the production process comes from raw materials, especially coke ash and gangue from the metallic charge. In order to improve the quality of the final product, it is necessary to use raw materials with low variations in chemical composition and with low silicon content, and to keep it as constant as possible with respect to its optimum level, in order to minimize the costs of secondary refining in the steelworks' converters. It is also worth noting that the

excess silicon in the hot metal requires a greater amount of calcium oxide (CaO) in the steel mill to perform secondary refining, resulting in a greater amount of slag and increasing production costs<sup>2,10,16</sup>.

In this sense, silicon content prediction models are useful for the production process. They support the reactor operation and allow to work with smaller safety margins, to optimize the fuel consumption and to improve the reactor efficiency.

In conducting a comparison between machine learning and hot metal metallurgy and evaluating the synaptic weights of the models, it is found that the most important variables are sinter and blowing flow. As for sinter, it contains SiO<sub>2</sub>, which serves as a Si source for hot metal, which could explain its influence on the model. Regarding the blowing flow, higher values favor a stronger blowing penetration and affect the thermal level of the blast furnace, which affects the silicon content incorporated into hot metal.

Other variables that also have an effect on silicon content were the enrichment of O<sub>2</sub>, pressure, and the amount of air blow in the tuyeres. These are variables that can affect the shape, thickness and position of the cohesive zone and consequently the silicon content of the hot metal.

For example, low gas permeability may indicate a thicker cohesive zone. Increasing the O<sub>2</sub> enrichment tends to increase the reactor temperature and decrease the amount of nitrogen injected, increasing the thermal level and favoring the permeability of the blast furnace. Thus, a high or thicker cohesive zone causes an increase in the silicon content in the hot metal<sup>10</sup>.

Finally, it turns out that the silicon introduced into the process is dissipated into the hot metal and slag. Thus, based on the binary basicity (CaO/SiO<sub>2</sub>) and using a mass balance calculation, it is possible to determine the silicon content in the hot metal. This information underlines the importance of controlling the conditions that affect the basicity of the slag. It is therefore entirely justifiable to use artificial neural networks to predict the silicon content during hot metal production<sup>10,16-18</sup>.

## 4. Conclusions

The increasing development of computing capacity, leading to cheaper devices with greater capacity, has driven the development of more complex algorithms with better results, as is the case with neural networks;

It should be noted that the most important part of modeling a neural network is the previous treatment of the database to be used for model development;

The neural network is an interesting tool to support decision making and operational planning in terms of fuel economy, operational stability, and delivery of a quality product for the steelworks, and helps to improve process monitoring;

Neural networks are tools capable of predicting silicon content based on parameters of the reduction process in blast furnaces and this can be verified by the precision of the model;

The Pearson and MSE correlation coefficient values confirmed that the hidden layer with 30 neurons gave the best results;

The analysis of the synaptic weights confirmed that the blower air, the sinter, the oxygenation and the pressure have a greater influence on the silicon content in the hot metal;

In order to a lesser extent, the slag rate, ( $\text{SiO}_2$ ) and ( $\text{CaO}$ ) have a lesser influence on the variation of silicon content and do not directly contribute to this silicon to hot metal transfer mechanism;

In conclusion, neural networks can be used in practise due to the excellent correlations between real values and the values calculated by the neural network.

## 5. References

1. Arif MS, Ahmad I. Artificial intelligence based prediction of exergetic efficiency of a blast furnace. *Computer-Aided Chem Eng.* 2021;50:1047-52.
2. Cardoso W, di Felice R, Baptista RC. Mathematical modeling to predict fuel consumption in a blast furnace using artificial neural networks. In: Márquez FPG, editor. *International Conference on Intelligent Emerging Methods of Artificial Intelligence & Cloud Computing: Proceedings of IEMAICLOUD 2021*. London, UK: Springer; 2021. In press.
3. Chen J. A predictive system for blast furnaces by integrating a neural network with qualitative analysis. *Eng Appl Artif Intell.* 2001;14(1):77-85.
4. Chizhikova VM. Best available techniques in the blast-furnace production. *Metallurgist.* 2020;64(1-2):13-35.
5. Li J, Hua C, Qian J, Guan X. Low-rank based Multi-Input Multi-Output Takagi-Sugeno fuzzy modeling for prediction of molten iron quality in blast furnace. *Fuzzy Sets Syst.* 2021;421:178-92.
6. Matino I, Dettori S, Colla V, Weber V, Salame S. Two innovative modelling approaches in order to forecast consumption of blast furnace gas by hot blast stoves. *Energy Procedia.* 2019;158:4043-8.
7. Fontes DOL, Vasconcelos LGS, Brito RP. 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.* 2020;141:107028.
8. Pavlov AV, Polinov AA, Spirin NA, Onorin OP, Lavrov VV. Use of model systems for solving new technological problems in blast-furnace production. *Metallurgist.* 2017;61(5-6):448-54.
9. Pettersson F, Chakraborti N, Saxén H. A genetic algorithms based multi-objective neural net applied to noisy blast furnace data. *Appl Soft Comput.* 2007;7(1):387-97.
10. Radhakrishnan VR, Mohamed AR. Neural networks for the identification and control of blast furnace hot metal quality. *J Process Contr.* 2000;10(6):509-24.
11. Chen M, Wan X, Shi J, Taskinen P, Jokilaakso A. Experimental study on the phase relations of the  $\text{SiO}_2$ -MgO-TiO<sub>2</sub> system in air at 1500°C. *JOM.* 2021. In press.
12. Dučić N, Jovičić A, Manasijević S, Radiša R, Čojbašić Ž, Savković B. Application of machine learning in the control of metal melting production process. *Appl Sci.* 2020;10(17):6048.
13. He F, Zhang L. Prediction model of end-point phosphorus content in BOF steelmaking process based on PCA and BP neural network. *J Process Contr.* 2018;66:51-8.
14. Hou Y, Wu Y, Liu Z, Han H, Wang P. Dynamic multi-objective differential evolution algorithm based on the information of evolution progress. *Sci China Technol Sci.* 2021;64(8):1676-89.
15. Pavlov AV, Polinov AA, Spirin NA, Onorin OP, Lavrov VV, Gurin IA. Decision-making support in blast-furnace operation. *Steel Transl.* 2019;49(3): 185-93.
16. David SF, David FF, Machado MLP. Artificial neural network model for predict of silicon content in hot metal blast furnace. *Mater Sci Forum.* 2016;889(1):572-7.
17. Diniz APM, Côco KF, Fomes FSV, Salles JLF. Forecasting model of silicon content in molten iron using wavelet decomposition and artificial neural networks. *Metals.* 2021;11(7):1001.
18. Dobrzanski LA, Trzaska J, Gawron M, Berliński M. The use of artificial neural networks for the prediction of sulphur content in hot metal produced in blast furnace. *Journal of Achievements in Materials and Manufacturing Engineering.* 2015;70(2):86-92.
19. Saxén H, Pettersson F. Nonlinear prediction of the hot metal silicon content in the blast furnace. *ISIJ Int.* 2007;47(12):1732-7.
20. Liu Y, Wang Y, Chen L, Zhao J, Wang W, Liu Q. Incremental Bayesian broad learning system and its industrial application. *Artif Intell Rev.* 2021;54(5):3517-37.
21. Jantre SR, Bhattacharya S, Maiti T. Quantile regression neural networks: a bayesian approach. *J Stat Theory Pract.* 2021;15(3):68.
22. Kina C, Turk K, Atalay E, Donmez I, Tanyildizi H. Comparison of extreme learning machine and deep learning model in the estimation of the fresh properties of hybrid fiber-reinforced SCC. *Neural Comput Appl.* 2021;33(18): 11641-59.
23. Liang W, Wang G, Ning X, Zhang J, Li Y, Jiang C, et al. Application of BP neural network to the prediction of coal ash melting characteristic temperature. *Fuel.* 2020;260:116324.
24. Reynolds QG, Rhamdhani MA. Computational modeling in pyrometallurgy: part I. *JOM.* 2021;73(9):2658-9.
25. Rhamdhani MA, Reynolds QG. Computational modelling in pyrometallurgy: part II. *JOM.* 2021;73(10):2885-7.
26. Zhang J, Li S, Li Z. Investigation the synergistic effects in quaternary binder containing red mud, blast furnace slag, steel slag and flue gas desulfurization gypsum based on artificial neural networks. *J Clean Prod.* 2020;273:122972.
27. Semenov YS, Gorupakha VV, Kuznetsov AM, Semion IY, Schumel'chik EI, Vashchenko SV, et al. Experience of using manganese-containing materials in blast-furnace charge. *Metallurgist.* 2020;63(9-10):1013-23.
28. Stein S, Leng C, Thornton S, Michel R. A guided analytics tool for feature selection in steel manufacturing with an application to blast furnace top gas efficiency. *Comput Mater Sci.* 2021;186:110053.
29. Zhang X, Kano M, Matsuzaki S. A comparative study of deep and shallow predictive techniques for hot metal temperature prediction in blast furnace ironmaking. *Comput Chem Eng.* 2019;130:106575.
30. Xie J, Zhou P. Robust stochastic configuration network multi-output modeling of molten iron quality in blast furnace ironmaking. *Neurocomputing.* 2020;387:139-49.
31. Völker C, Firdous R, Stephan D, Kruschwitz S. Sequential learning to accelerate discovery of alkali-activated binders. *J Mater Sci.* 2021;56(28):15859-81.
32. Xu Z, Zheng Z, Gao X. Operation optimization of the steel manufacturing process: a brief review. *Int J Miner Metall Mater.* 2021;28(8):1-21.
33. Zhan ZH, Shi L, Tan KC, Zhang J. A survey on evolutionary computation for complex continuous optimization. *Artif Intell Rev.* 2021. In press.