

Article

# Machine Learning-Based Prediction of a BOS Reactor Performance from Operating Parameters

Alireza Rahnama <sup>1,\*</sup>, Zushu Li <sup>2</sup> and Seetharaman Sridhar <sup>1,3</sup>

<sup>1</sup> AI-GARISMO, 1 Sandover House, 124 Spa Road, London SE16 3FD, UK; sseetharaman@mines.edu or sridhar.seetharaman@ai-garismo.com

<sup>2</sup> WMG, The University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK; z.li.19@warwick.ac.uk

<sup>3</sup> George S. Ansell Department of Metallurgical and Materials Engineering, Colorado School of Mines, Golden, CO 80401, USA

\* Correspondence: alireza.rahnama@ai-garismo.com

Received: 25 November 2019; Accepted: 13 March 2020; Published: 23 March 2020



**Abstract:** A machine learning-based analysis was applied to process data obtained from a Basic Oxygen Steelmaking (BOS) pilot plant. The first purpose was to identify correlations between operating parameters and reactor performance, defined as rate of decarburization (dc/dt). Correlation analysis showed, as expected a strong positive correlation between the rate of decarburization (dc/dt) and total oxygen flow. On the other hand, the decarburization rate exhibited a negative correlation with lance height. Less obviously, the decarburization rate, also showed a positive correlation with temperature of the waste gas and CO<sub>2</sub> content in the waste gas. The second purpose was to train the pilot-plant dataset and develop a neural network based regression to predict the decarburization rate. This was used to predict the decarburization rate in a BOS furnace in an actual manufacturing plant based on lance height and total oxygen flow. The performance was satisfactory with a coefficient of determination of 0.98, confirming that the trained model can adequately predict the variation in the decarburization rate (dc/dt) within BOS reactors.

**Keywords:** machine learning; artificial intelligence; neural network; BOS reactor; steelmaking

## 1. Introduction

The processing of lower grade ores is a topic of particular interest, as fluctuation in raw material cost is a key challenge to sustainability in the steel industry. Raw materials flexibility is to a great extent enabled in the basic oxygen steelmaking (BOS) process, wherein oxidizable impurities, such as phosphorus (P) and silicon (Si), are separated into a slag phase. However, the primary purpose of the BOS process is to convert pig iron into crude steel and, therefore, impurity removal being a secondary function of the converter, has to be balanced by the decarburization process.

As a result, processing parameters have to be balanced, and many different chemical reactions compete for the available oxygen. There are also many processing parameters that are interconnected in complicated ways that cannot be readily predicted. The BOS process is in general advised by static mass- and heat-balanced models, which predict, based on inputs, the resultant end-point. The input parameters include at least the quantities of hot metal, steel scrap, iron ore, fluxes and oxygen to blow, while the resultant end-point is the temperature, weight and composition of steel produced. The process control is supported by various measurements, such as in-blow sampling, end-blow sampling, oxygen flow rate/ total oxygen blown, lance height, waste gas flow rate, waste gas pressure, waste gas composition, etc. The operating information is recorded and stored in a plant manufacturing system normally at the interval of one second. Dynamic process modeling of the BOS process, depending on the researchers, can be based on thermodynamics [1], multizone reactions [2–4], empirical or

semi-empirical plant data regression, in combination with computational fluid dynamics (CFD) in some cases [5]. Data-driven modeling techniques have been applied in the dynamic modeling of basic oxygen processes. These include the artificial neural network (ANN) [6–8], fuzzy logic [9], support vector machine [10,11] and case base reasoning [12]. However, most of these models are limited to the end-point control of carbon and temperature.

The pilot plant work published by Millman et al. [2] (“IMPHOS” or Improved phosphorus refining) provides a set of data on top-blown oxygen steelmaking under conditions similar to an actual BOS plant. The trials were carried out in a six-ton pilot converter at the MEFOS in Sweden and provides data for heats through samples obtained from seven vertical positions in the vessel at 2-min intervals. In addition, 28 process data parameters (including decarburization rate) were recorded every second. The major parameters recorded are the total of materials added as a function of time (hot metal and fluxes); oxygen flow rate/total amount of oxygen blown (total oxygen flow); lance height; waste gas (temperature; flow rate and composition of CO, CO<sub>2</sub>, N<sub>2</sub> and O<sub>2</sub>) and the data calculated from the above measurements, for example,  $dc/dt$  (the decarburization rate).

Using machine learning (ML) approaches to predict material properties is a research field that has been increasing in the last years [13]. Thermodynamic stability of perovskite, transition temperatures of inorganic glasses, glass-forming ability of metallic alloys and interphase precipitation in HSLA (High Strength Low Alloy) steels, among others, are examples of material properties studied by machining learning approaches [14–18]. Machine learning algorithms enable us to design new grades of alloys purely based on past data, either available as an open dataset or reported in the scientific literature [19].

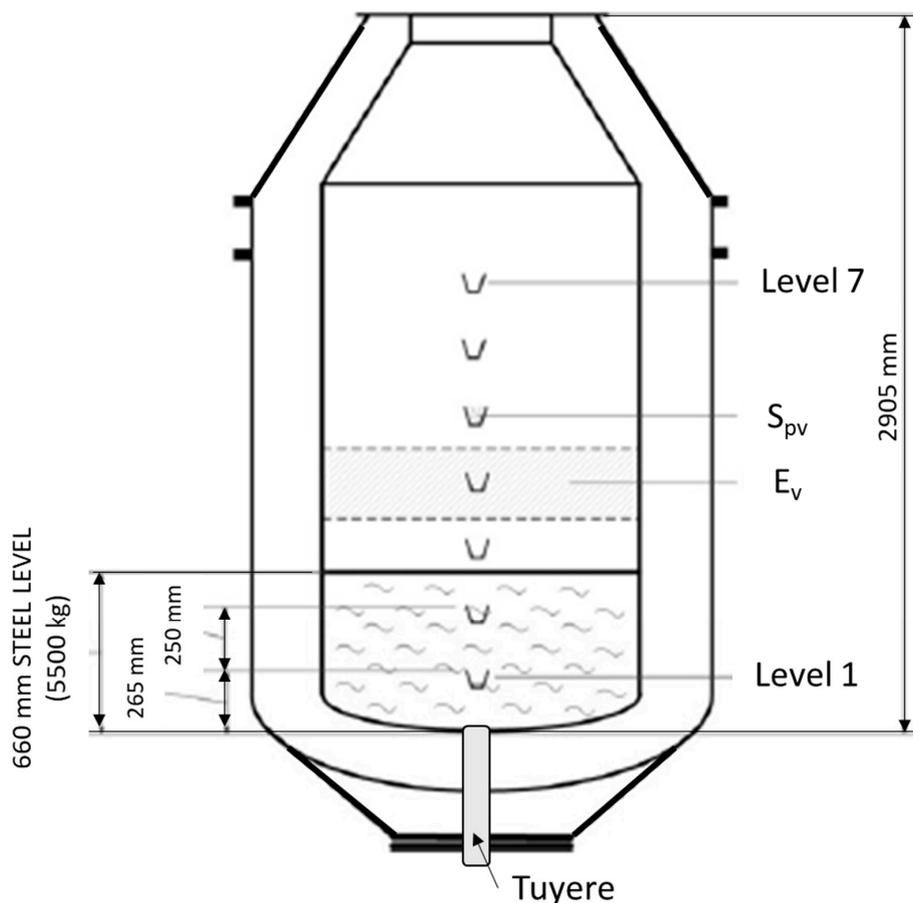
The objective of this study is to carry out a machine learning (ML) analysis on the data generated from the six-ton pilot BOS trials (IMPHOS data) with the purposes of (i) identifying correlations between features and (ii) based on a training data and test data developing a neural network-based regression model (algorithm) to predict the decarburization rate. The tools that we employed in this study were Microsoft Power BI for statistical analysis; Python (Scikit-learn, Matplotlib and Seaborn) for performing correlation analysis and statistical analysis and Microsoft Azure Machine Learning Studio for regression models and their subsequent performance evaluation [20]. The algorithm developed here is used to successfully predict the decarburization rate ( $dc/dt$ ) in an actual manufacturing plant based on two operating parameters of lance height and total oxygen flow.

## 2. The Multiphysics of the Basic Oxygen Steelmaking Process

The overall purpose of the BOS process is to refine hot metal (pig iron) to a steel of desired chemistry (C, Si, Mn, P and S) and temperature in a controlled manner in a refractory converter (as shown schematically in Figure 1). In the BOS process wherein C-rich pig iron is converted to steel, pure oxygen is injected through a water-cooled lance into the molten metal bath, and the bath is covered by molten oxides (slag phase). The role of the oxygen is to selectively react with oxidizable elements and primarily remove the C (from >4% to less than 0.1%) as a gas phase but also remove other unwanted impurities, stemming from the iron ore as oxides, that separate and dissolve into the slag phase.

The kinetics of the BOS process is very complex, because it involves multiple simultaneous processes. Simultaneous multiphase interactions, heat and mass transfer, gas-slag-metal chemical reactions in multiple zones and vigorous fluid flow caused by the impingement of the oxygen jet occur in a BOS reactor at high temperatures. In addition, it is a dynamic and transient process, which makes the kinetics involved in a reactor more complex. Direct measurements of temperatures and chemistries are very difficult due to the nature of the process, which involves harsh conditions. That is why many researchers have been trying to address these difficulties through modeling the process.

A multiphysics description of the converter process, with the ultimate aim of predicting carbon content in the melt, will have to involve several sub-models at different scales. These models will have to capture various phenomenon occurring in the gas/metal, metal/slag and metal/slag/gas mixtures, as well as transport processes in the bulk metal and bulk slag. The models include, but are not limited to, the following:



**Figure 1.** A schematic of the pilot plant BOF converter used in the Improved Phosphorus-refining (IMPHOS) trials with dimensions and indicated location heights of the sampling levels [2].

- (i) Multiphase fluid flow in the various multiphase mixtures. Since this is a complex and computational task by itself, it is often simplified to 2D and informed by physical water models [5,21,22].
- (ii) Reaction kinetics and competition between the injected oxygen and various dissolved elements, such as C, but also P, Si, Mn, etc. [4,23].
- (iii) Microscale interactions between the phases, which include the separation of elements and precipitated phases across interfaces and emulsification [4,24,25].

The decarburization process mainly takes place in two reaction zones: namely, the jet impact zone and the gas-slag-metal emulsion zone [4,24]. At the emulsion zone, a sequence of chemical reactions take place: Firstly, FeO reacts with CO to form CO<sub>2</sub> at the slag-gas interface. Secondly, CO<sub>2</sub> transfers from the slag-gas interface to the gas-steel droplet interface via the gas phase. Subsequently, CO<sub>2</sub> reacts with carbon in an iron droplet at the metal-gas interface, and the formed CO is transferred back to the slag-gas interface via the gas phase. It should be noted that there is disagreement among researchers about the description of the mechanisms and the rate-controlling step(s) under different operating conditions [26]. Some other studies proposed that the reaction rate is controlled by the interfacial area change because of material exchanges during the process. The phase field models developed according to the interfacial instability normally consider only the interfacial tension and fluid flow [24]. The coupling of models across scales is a hard enough (and computationally intensive) problem, which can be simplified and described under simplified process windows. However, the predictive capability under widely ranging process parameters is not possible. Consequently, all the previous studies, although mechanistic in nature, have different levels of simplification to address the complex phenomena in the BOS reactor.

The removal of elements (C, Si, Mn, P and S) from pig iron in the industrial BOS process is mainly controlled by two operating parameters: lance height (i.e., the height of the lance tip above the static steel bath) and oxygen blowing rate (or total oxygen flow/blown). The control strategy can be varying one or both of these two parameters. The lance profile (i.e., lance height as a function of blowing time) is critically important for the efficient production of high-quality steel. For example, a high lance position can help stir and oxidize the slag with a high FeO content, which can accelerate the removal of carbon. A low lance height can substantially increase the impingement of the oxygen jet into the metal bath and splash the metal droplets into the slag layer, which accelerates the formation of the emulsion zone and the removal of carbon and other elements through the interfacial reactions between the massive numbers of metal droplets with emulsified slag [27]. The refining behavior (e.g., the decarburization rate) has a close relationship with the lance profile and oxygen blowing rate (or total oxygen flow). Thus, in this work, we intend to develop a novel algorithm to predict the refining behavior of the BOS reactor (using the decarburization rate as an example) by using machine learning to analyze the massive dataset of operating parameters, including lance height and total oxygen flow.

Using machine learning, we propose a technique that does not employ any simplifying assumptions. On the contrary, the algorithm was trained on a real dataset. In addition, the machine learning algorithm does not require to take into account any of the physics involved in the process. It actually provides a ring road to all the complexities involved in a BOS reactor. In addition, as shown later, the technique is able to predict the decarburization rate precisely. The aim of the current study is, thus, twofold: (1) Firstly, to identify the correlations and trends from a set of processing parameters on carbon removal. (2) Secondly, to try and circumvent the granularity of multiphysics and develop an artificial intelligence-based predictive model for carbon removal.

### 3. Dataset

A six-ton BOF (Basic Oxygen Furnace) pilot plant converter located at Swerea MEFOS, Sweden, as part of the European commission funded project IMPHOS: Improved Phosphorus refining [2]. In this reactor, operating conditions, such as oxygen blow rate (total oxygen flow), lance height, the quantities of hot metal and fluxes used and off-gas analysis, are recorded. Gas/slag/metal emulsion samples were taken from seven various heights and at 2-min intervals from the start of blowing during a blow via robotic delivery. The sampler lances consist of an inner and outer structure of three mild steel bars, the inner being joined to the sides of inline sample pots and the outer joined to the sides of the disc sample pot lids. The sampler lance is lowered into the converter through an opening in the top (with a slight offset from the oxygen lance), with the lids in the closed position. Once lowered into the position, the outer three bars are retracted, lifting the lids in situ and allowing the sample pots to fill. A schematic of the converter is given in Figure 1 with dimensions. The converter had a lining of magnesia-carbon (fired and fused) bricks. Gas inlets include a single bath agitation tuyere, blowing nitrogen at a rate of  $0.5 \text{ Nm}^3 \text{ min}^{-1}$ , and a water-cooled oxygen lance blowing oxygen at a nominal rate of  $17 \text{ Nm}^3 \text{ min}^{-1}$ . During the blow, the following was measured as output data from the reactor: oxygen flow rate/total oxygen flow; lance height and waste gas (temperature; flow rate and composition of CO, CO<sub>2</sub>, N<sub>2</sub> and O<sub>2</sub>).

There are various models, such as static (mass balance model, thermodynamic model, etc.); empirical (regression based on individual plant data) and dynamic. These models are intended to predict the end-point of steel or the steel/slag compositions as a function of blowing time and to control/optimize BOF steelmaking. In the BOF process, the blown oxygen is used for decarburization; the removal of other impurities of Si, Mn and P; the formation of iron oxides in the slag and post-combustion. A typical example is off-gas information. The industry has been exploring the use of off-gas information for process control. The most important one is to use the off-gas composition to monitor the decarburization rate in the BOS converter. If the decarburization rate can be precisely controlled or predicted, then under the known oxygen-blowing conditions, the amount of oxygen in the slag (iron oxides) and the amount of oxygen to remove other elements, especially for phosphorus,

can be predicted. The decarburization rate can be calculated by the off-gas analysis (CO, CO<sub>2</sub>, N<sub>2</sub> and O<sub>2</sub>). Then, we will link the decarburization rate (calculated from off-gas composition) to the operating parameter(s). This will provide the foundation to predict the decarburization rate (dc/dt) and, as a result, dynamically control the converter operation. The features we are investigating from the dataset are presented in Table 1. The other operating conditions are listed in Table 2, including the chemistries of hot metal; steel and slag; temperature; charges (hot metal, scrap and fluxes) and lance height.

**Table 1.** The operating parameters and their features obtained from the six-ton pilot plant BOF converter system.

Parameter	Feature
Charge Time (min)	Blowing time (in minutes from O <sub>2</sub> ignition)
dc/dt (kg C/min)	Instant decarburization rate
Oxygen Yield (%)	Instant oxygen yield
Total Oxygen flow rate (Nm <sup>3</sup> /min)	Instant oxygen flow rate
Total C removal (kg)	Calculated total C removed from the start of O <sub>2</sub>
Total N <sub>2</sub> flow rate (Nm <sup>3</sup> /min)	Instant N <sub>2</sub> flow rate
Total O <sub>2</sub> (Nm <sup>3</sup> /min)	Total blown oxygen volume
Total N <sub>2</sub> (Nm <sup>3</sup> /min)	Total bottom blown nitrogen volume
Total propane (Nm <sup>3</sup> /min)	Total blown propane volume
Lance height (mm)	Instant lance height from the point of calculated metal bath level
dC/dt (kg C/s)	Calculated from off-gas composition
dO/dt (kg O/s)	Calculated from off-gas, oxygen for decarburization
dO <sub>s</sub> /dt (kg O/s)	Calculated from off-gas composition, oxygen into slag
Temp. waste gas (°C)	Off-gas temperature (measured after water cooling)
CO waste gas (%)	Measured waste gas composition
CO <sub>2</sub> waste gas (%)	Measured waste gas composition
O <sub>2</sub> waste gas (%)	Measured waste gas composition

Note: dC/dt (kg C/s) is calculated from the waste gas composition, which was recorded in the operating system in every second, while dc/dt (kg C/min) is converted from dC/dt (kg C/s). The decarburization rate is generally expressed as dc/dt (kg C/min), and therefore, the target of this study is to predict dc/dt (kg C/min) by a machine-learning (ML)-based algorithm from the operating parameters. It should be pointed out that dC/dt (kg C/s) is still mentioned in this study for the purpose of explanation.

In order to validate the correlation between the dc/dt and the operating parameters (the ML-based methodology, as explained in Section 4, Methods, to predict dc/dt from operating parameters) established from the six-ton BOF converter, production information has been taken from a 330-ton converter at the Tata Steel UK Port Talbot Plant. The operating parameters and the resultant information include oxygen flow rate; total oxygen flow; off-gas analysis (temperature, rate, composition CO, CO<sub>2</sub>, N<sub>2</sub> and O<sub>2</sub>) and dC/dt (kg C/s) calculated from the off-gas analysis. A total of 1100 data points were obtained at different times for each value corresponding to Table 1, from the 6-ton pilot reactor, which were used for training and test data. In addition, a total of 1200 points corresponding to known conditions in the 330-ton full scale reactor were used as additional test data to investigate if the trained model could predict the decarburization rate in a scaled-up system. It should be noted that differences in the levels/quantities of the parameters exist between the pilot and industrial converters; for example, the heat size (6t for pilot converter and 330t for industrial converter), blowing time (~18 min for the 6t pilot converter and ~20 min for the 330t industrial converter), lance height (110~180 mm for the 6t pilot converter and 2.0~2.6 meters for the 330t industrial converter), O<sub>2</sub> flow rate (14.00~17.30 Nm<sup>3</sup>/min for the 6t pilot converter and 598~1025 Nm<sup>3</sup>/min for the 330t industrial converter) and total oxygen

flow (278 Nm<sup>3</sup> for the 6t pilot converter and 18863 Nm<sup>3</sup> for the 330t industrial converter). However, the parameters and their features recorded in their systems are similar for both conditions, and in this paper, we use the dataset from one heat of both the 6t pilot and the 330t industrial converters, respectively, to develop and demonstrate the machining learning-based model for the prediction of the decarburization rate from operating parameters under different conditions.

**Table 2.** The operating conditions for both the six-ton (6t) and 330t converters.

6t pilot BOF	C	Si	Mn	P	S	T (°C)	
Hot metal (%)	3.78~4.25	0.41~0.88	0.39~0.48	0.067~0.095	0.037~0.081	1272~1316	
Steel (%)	0.01~0.41	0~0.15	0.05~0.27	0.008~0.042	0.018~0.035	1669~1772	
	CaO	SiO <sub>2</sub>	FeO	MnO	Al <sub>2</sub> O <sub>3</sub>	MgO	P <sub>2</sub> O <sub>5</sub>
Slag (%)	26.2~52.3	6.7~19.2	8.3~30.5	2.6~4.1	0.7~1.6	4.6~17.7	0.73~1.62
330t BOF	C	Si	Mn	P	S	T (°C)	
Hot metal (%)	3.60~5.18	0.30~1.20	0.11~0.52	0.058~0.125	0.025~0.072	1341~1412	
Steel (%)	0.021~0.55	0~0.10	0.03~0.21	0.001~0.057	0.011~0.032	1579~1738	
	CaO	SiO <sub>2</sub>	FeO	MnO	Al <sub>2</sub> O <sub>3</sub>	MgO	P <sub>2</sub> O <sub>5</sub>
Slag (%)	36.07~55.35	9.13~19.58	12.15~31.34	2.04~5.62	0.64~4.97	3.61~10.73	0.94~2.38
	6t pilot BOF		330t BOF				
hot metal	4410~5170 kg		269.6~302.5 t				
scrap	500~750 kg		46.0~85.5 t				
lime	250~350 kg		5~25 t				
dolomet	0~30 kg		0~11.5 t				
iron ore	0 kg		0~6.5 t				
total oxygen	263~307 Nm <sup>3</sup>		12265~21641 Nm <sup>3</sup>				
Lance height	110~180 mm		2.0~2.6 m				

#### 4. Method

While neural networks are mainly known for applications in deep learning, they can be easily used for regression problems and are especially suitable when linear regression does not work. Neural network regression is another supervised learning algorithm available in Microsoft Azure Machine Learning Studio and requires a label column which has to be numerical [20]. Since our label column (dc/dt) was numerical, and the results generated by classical linear regression had low accuracy, we used a neural network regression. Linear regression resulted in a coefficient of determination of 0.45 (Table 3, column 2), which was much lower than that by using neural network regression (0.99) and deemed to be below what is acceptable. The output of a single neuron has a form of  $g(\sum_j w_{ji}x_j)$ , where  $w_{ji}$  are the weights and  $x_j$  are the input. A continuous activation function, e.g., sigmoid function of a form of  $1/(1 + e^{-x})$ , was employed. We used “Parameter Range” in the Create trainer mode and subsequently employed the Tune Model Hyperparameters module to iterate over the possible combinations of parameters to achieve the optimal configuration. In “Hidden layer specification” we selected “fully connected case” to create a model that has exactly one hidden layer, and the output layer is fully connected to the hidden layer, and the hidden layer is fully connected to the input layer. Number of nodes in the hidden layer and learning rate were used as the hyperparameters. The optimized number of hidden nodes was calculated to be 100. Another hyperparameter that was tuned was the learning rate. Learning rate is the step taken at each iteration before correction. A large value of the learning rate makes the convergence faster; however, it can result in overshooting the local minima. The tuned learning rate was calculated to be 0.1. Min-max normalizer was selected to linearly rescale each feature to the [0, 1] interval. Rescaling to the [0, 1] interval was carried out by shifting the values of each features, i.e., the minimum value is 0, and then dividing by the new maximum value. The error after each iteration is calculated and “back-propagated” to the network using the chain rule.

**Table 3.** Evaluation metrics of the predicted values of the  $dc/dt$  by using different features in the pilot dataset and using different regression methods.

Evaluated Metrics	Pilot Dataset with All Features <sup>(1)</sup>	Pilot Dataset with All Features <sup>(2)</sup>	Pilot Dataset without $dc/dt$ <sup>(2)</sup>	Pilot Dataset with Total O <sub>2</sub> Flow and Lance Height Only <sup>(2)</sup>	Industrial Dataset with Total O <sub>2</sub> Flow and Lance Height Only <sup>(2)</sup>
Mean Absolute Error	0.12	0.029	0.030	0.034	0.25
Root Mean Square Error	0.51	0.043	0.055	0.060	0.62
Relative Absolute Error	0.42	0.005	0.006	0.008	0.04
Relative Square Error	0.48	0.000046	0.00006	0.0001	0.009
Coefficient of Determination	0.45	0.99	0.99	0.97	0.98

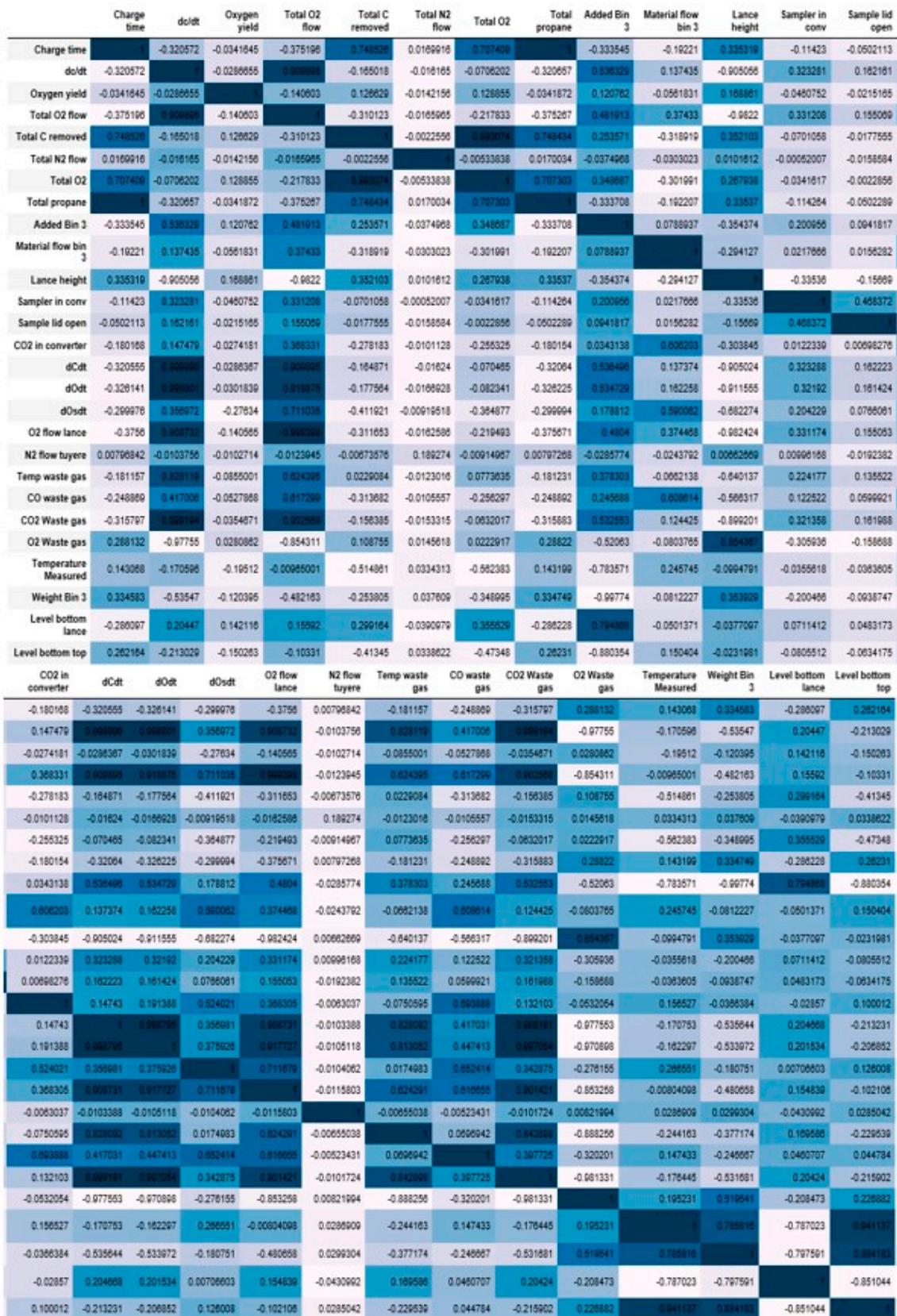
(1) Using linear regression and (2) using neural network regression.

## 5. Results and Discussion

### 5.1. Machine Learning Predictions of the Decarburization Rate $dc/dt$

Figure 2 shows the correlation matrix between different features, which was generated by the available dataset. The correlation between two datasets varies between -1 and 1, with 0 implying no correlation. Correlation of +1 implies a perfect positive correlation (e.g., as  $x$  increases, so does  $y$ ), and -1 implies a perfect negative correlation (as  $x$  increases,  $y$  decreases). Dark blue implies a strong positive correlation, and lighter pink shows a strong negative correlation. For example, there is a strong positive correlation between Total O<sub>2</sub> flow and  $dO/dt$ , or there is a strong negative correlation between O<sub>2</sub> waste gas and  $dc/dt$ .

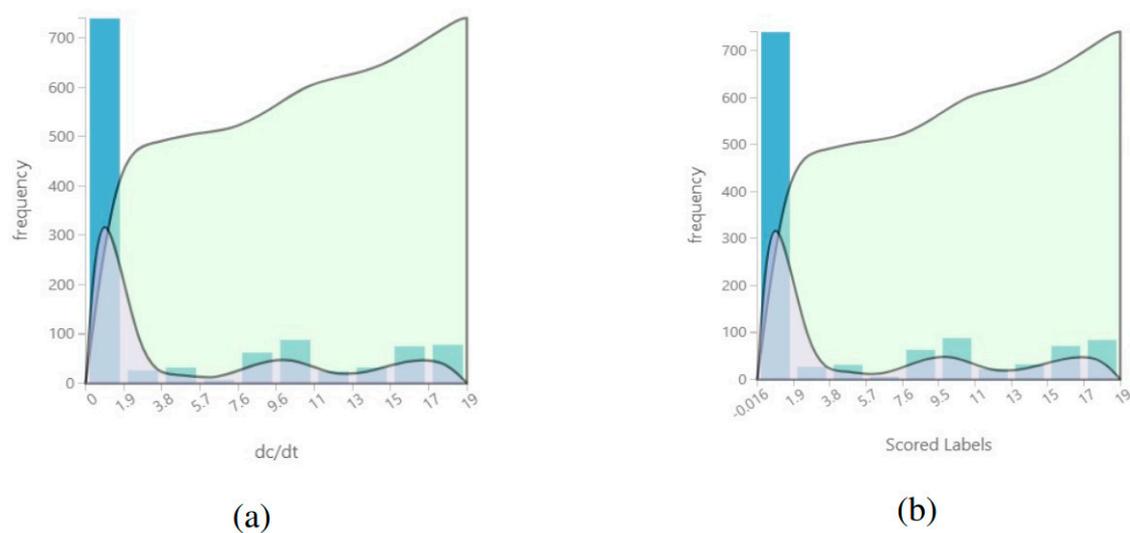
Total oxygen flow has nearly perfect positive ( $>0.9$ ) correlation with  $dc/dt$  (kg C/min) or  $dC/dt$  (kg C/s), total C removed,  $dO/dt$ ,  $dOs/dt$  (0.71) and waste gas CO<sub>2</sub> composition. The total oxygen flow was mainly used for decarburization escaping from the reactor as waste gas and remaining in the slag (oxidization of Si, Mn, Fe and P elements in the liquid metal). The former is directly linked with parameters  $dc/dt$  (or  $dC/dt$ ), total C removed,  $dO/dt$  and waste gas CO<sub>2</sub> composition, while the latter is in the form of  $dOs/dt$ . Furthermore,  $dc/dt$  (or  $dC/dt$ ) has nearly perfect positive correlation with total O<sub>2</sub> flow and waste gas CO<sub>2</sub> composition and nearly perfect negative correlation with lance height and waste gas O<sub>2</sub>. Except the oxygen from steel scrap and iron ore coolant, the main oxygen comes from the total oxygen blown through the lance that is also the main oxygen source for decarburization. Thus, the  $dc/dt$  (or  $dC/dt$ ) has a nearly perfect correlation with the total O<sub>2</sub> flow. During the pilot plant experiment, O<sub>2</sub> was blown at a fixed flow rate, and the refining performance in the converter was controlled by adjusting the lance height. The decarburization mainly occurred in two zones of a hot spot zone (at the vicinity of the location where the lance releases oxygen to the bath) and gas-slag-metal droplet emulsification zone (where the available area for slag/metal/gas reaction is high). Lower lance height increases the hot spot zone and the amount of metal droplets in the emulsification zone, and the latter increases the decarburization in the gas-slag-metal droplet zone. Therefore, the overall decarburization rate increases with decreasing the lance height, which explains the observed negative correlation. The decarburization rate is calculated from the waste gas composition according to the equation  $dc/dt \left( \frac{dc}{dt} = (CO + CO_2) \times waste\ gas\ flow\ rate \times \frac{12}{22.4} \right)$ . This explains well the perfect positive correlation between  $dc/dt$  (or  $dC/dt$ ) and waste gas CO<sub>2</sub> concentration. From the above analysis, both the total O<sub>2</sub> flow and the lance height are the controlling parameters for the decarburization, which indicates that the probability of predicting the decarburization rate by the combination of both parameters (see Section 5.3. Prediction of the  $dc/dt$  After Excluding Parameters). Finally,  $dO/dt$  has nearly perfect positive correlation with the total O<sub>2</sub> flow,  $dc/dt$  (or  $dC/dt$ ) and waste gas CO<sub>2</sub> composition but nearly perfect negative correlation with the lance height and waste gas O<sub>2</sub> content (similar explanation to that of the  $dc/dt$  or  $dC/dt$  dependence upon the parameters).



**Figure 2.** The correlations between different features within the dataset from the six-ton (6t) pilot plant trial. The figure has been vertically split into two parts, and the top and the bottom shown above are actually the left and right parts of the figure, respectively.

### 5.2. Prediction of $dc/dt$ With All the Features Included in the Dataset

Initially, we used all the features present in the dataset (including  $dc/dt$ ) to predict the value of the  $dc/dt$ . Figure 3 shows a statistical comparison between the actual values of the  $dc/dt$  (Figure 3a) and the predicted ones (Figure 3b). The blue bars show the histogram of the values. The green shadowed area shows the cumulative distribution function (CDF) and the blue shadowed area shows the probability density function (PDF). As shown in the figure, the predicted functions closely follow those of the actual values of the  $dc/dt$ . Both histograms show a frequency around 700 for a  $dc/dt$  value close to 0. In addition, both the histograms of the actual value of  $dc/dt$  and that of the predicted value show a bump in frequency around 100 for  $dc/dt$  values close to 9.6 and a bump at a frequency about 100 for  $dc/dt$  values around 17. As evident from Figure 3, the CDF of the actual values and that of the predicted values have a very similar shape. This confirms that the algorithm successfully captured the statistics of the dataset.

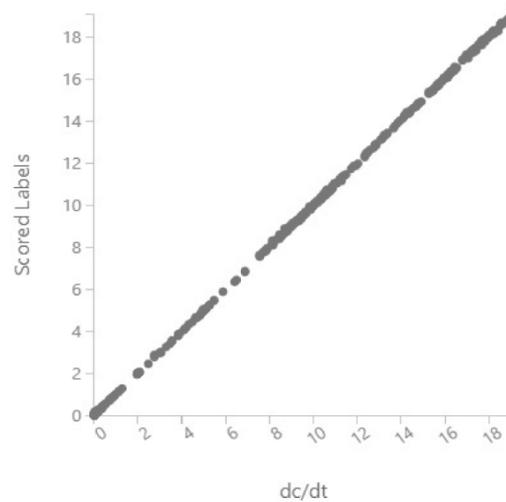


**Figure 3.** Histogram (blue), cumulative distribution function (green) and probability density function (purple) for (a) the actual values of the  $dc/dt$  and (b) the predicted values of the  $dc/dt$  for the pilot dataset with all the features included.

Table 4 compares the statistical metrics between the actual and predicted values of the  $dc/dt$ . Although the statistical metrics for the predicted values are slightly higher, it can be concluded that the machine-learning model successfully captured the statistics of the values of the  $dc/dt$ . This is further confirmed by Figure 4, the scatter plot, showing that the neural network regression model could accurately predict the values of the  $dc/dt$ . It should be noted that a scatter plot that compares the predicted values of the test set with the “true” values of the target is one the main metrics to evaluate a model performance. As is shown in Figure 4, the scored label as a function of the “true” value of the  $dc/dt$  follows a  $y = x$  line, meaning the model performed very well.

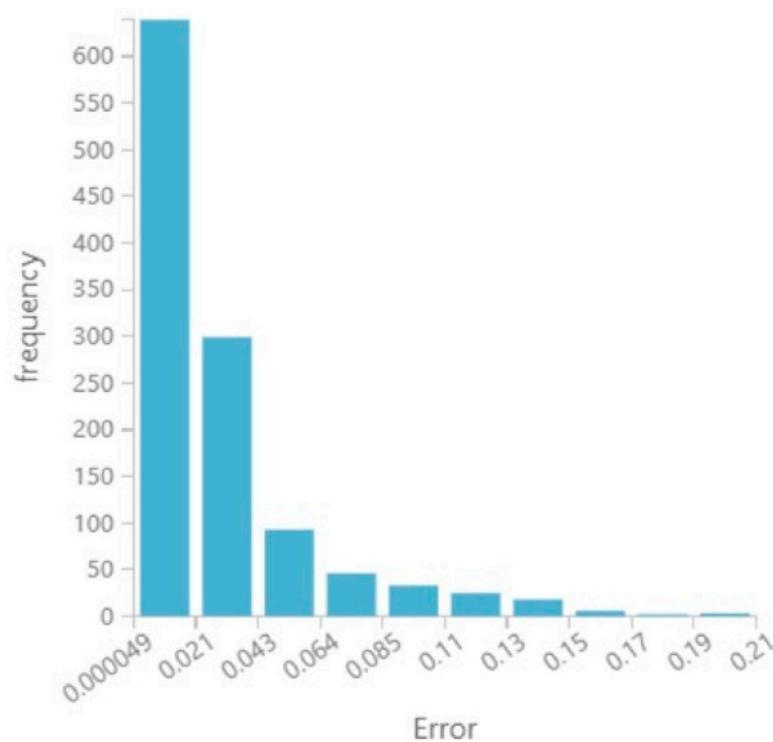
**Table 4.** Statistical comparison between the actual values of the  $dc/dt$  with the predicted values for the pilot dataset with all the features included.

	Mean	Median	Min	Max	Standard Deviation
Actual Statistics	4.43	0.081	0	19.105	6.42
Predicted Statistics	4.45	0.092	0	19.09	6.43



**Figure 4.** Scatter plot comparing the predicted values of the  $dc/dt$  using the neural network method with the actual values of the  $dc/dt$  for the pilot dataset with all the features included.

Figure 5 shows the error histogram of the neural network regression model. Errors with a value of 0.000049 had the highest frequency, confirming the excellent performance of the model. Table 3 (column 3) shows the performance metrics of the neural network regression model. The metrics were recorded to be 0.029 for the mean absolute error, 0.043 for the root mean squared error, 0.005 for the relative absolute error and 0.0046 for the relative squared error. The coefficient of determination for the model was calculated to be 0.99, which shows the excellent performance of the neural network regression algorithm.



**Figure 5.** Error histogram of the predicted values of the  $dc/dt$  by using all the features included in the pilot plant dataset.

We used the permutation method to measure the feature importance for the prediction of the  $dc/dt$ . The feature importance values are shown in Table 5. It was computed that the  $dC/dt$  had the highest importance, with a value of 9.06. In second place stands the  $dO/dt$  with a value of 0.16.

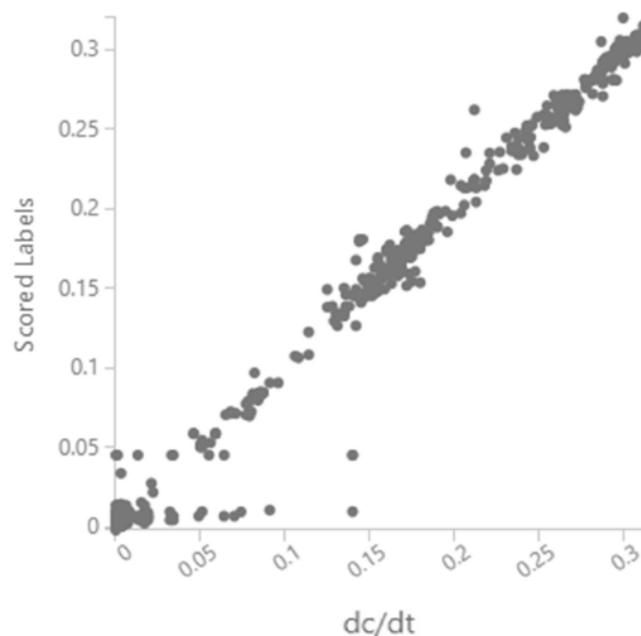
**Table 5.** Feature importance for predicting the values of the  $dc/dt$  in the pilot dataset with all the features.

$dC/dt$	$dO/dt$	Oxygen Yield	CO <sub>2</sub> Waste Gas	CO Waste Gas	Total O <sub>2</sub> Flow	$dO_s/dt$	Temp. of Waste Gas	Total O <sub>2</sub>	O <sub>2</sub> Waste Gas
9.06	0.16	0.07	0.04	0.004	0.002	0.0006	0.0002	0.00014	0.00013

### 5.3. Prediction of the $dc/dt$ After Excluding Parameters

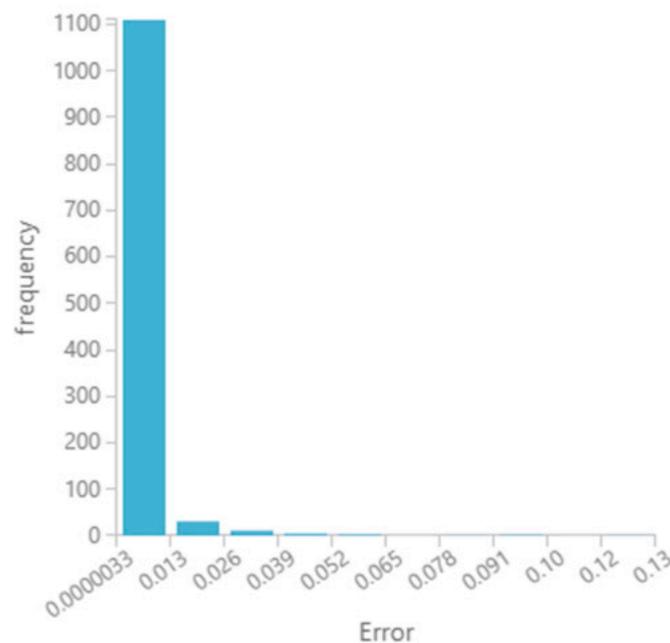
In order to establish whether the  $dc/dt$  could be predicted with reasonable accuracy with fewer parameters, parameters were successively removed. As discussed in the previous section, the  $dC/dt$  had a very high prediction power for the values of the  $dc/dt$ . However, it was anticipated that the  $dC/dt$  might lead to data leakage, as the  $dC/dt$  is calculated from the value of the  $dc/dt$ . Thus, the feature  $dC/dt$  was removed in order to measure the performance of the neural network regression model (Table 3, column 4). It was found that the scored labels (predicted values) have the very similar histogram, CDF and PDF graphs to those of the actual values of the  $dc/dt$ . Similar to the previous calculations, the statistical parameters belonging to the predicted values were slightly higher; however, the difference is very low, and it can be thus said that the model was able to capture the statistics of the data when the feature  $dC/dt$  was removed.

Then, we attempted to predict the value of the  $dc/dt$  using only two features: namely, total O<sub>2</sub> flow and lance height, because these are the two inputs that are controllable in an industrial reactor. The oxygen blown into the converter will be used for the decarburization ( $dO/dt$ ), oxidization of other elements into the slag ( $dO_s/dt$ ), oxygen in the waste gas, etc.; therefore, in the prediction of the  $dc/dt$ , the features of the  $dO/dt$ ,  $dO_s/dt$ , etc. are excluded. Figure 6 shows the scatter plot of the  $dc/dt$  versus the predicted values. As is evident from the figure, except for some values of the  $dc/dt$  where the predicted values were slightly lower, for most of the values, the neural network regression model was able to predict the  $dc/dt$  with a good accuracy.



**Figure 6.** Scatter plot comparing the predicted values of the  $dc/dt$  using the neural network method with the actual values of  $dc/dt$  by using the two features of total oxygen flow and lance height in the pilot dataset.

Figure 7 shows the error histogram for this prediction. The most frequent error was 0.0000033. Table 3 (column 5) shows the performance metrics for the prediction of the  $dc/dt$  using only two features. For this computation, the mean absolute error was calculated to be 0.034, root mean squared error to be 0.06, relative absolute error to be 0.008 and relative squared error to be 0.0001. The coefficient of determination was computed to be 0.97. These performance metrics showed that we were able to successfully predict the value of the  $dc/dt$  using only the two variables of total oxygen flow and lance height.



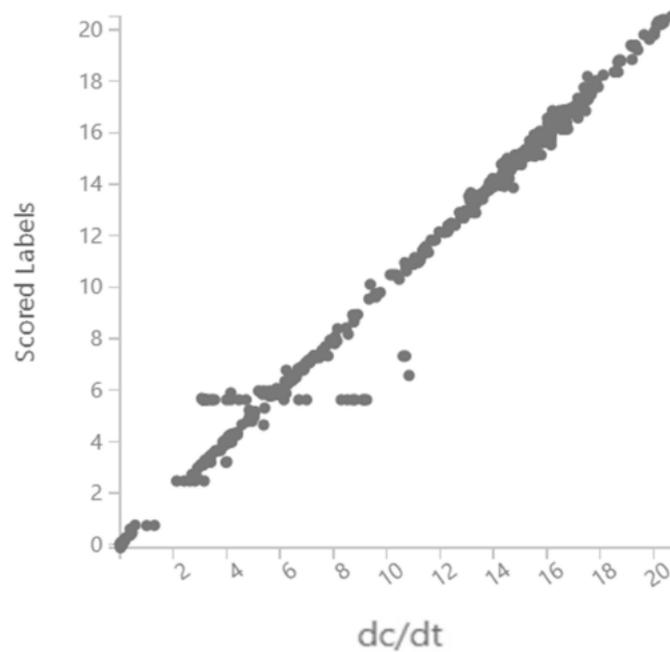
**Figure 7.** Error histogram of the predicted values of the  $dc/dt$  by using the two operating parameters of total oxygen flow and lance height in the pilot plant dataset.

#### 5.4. Prediction of the $dc/dt$ for an Industrial Dataset

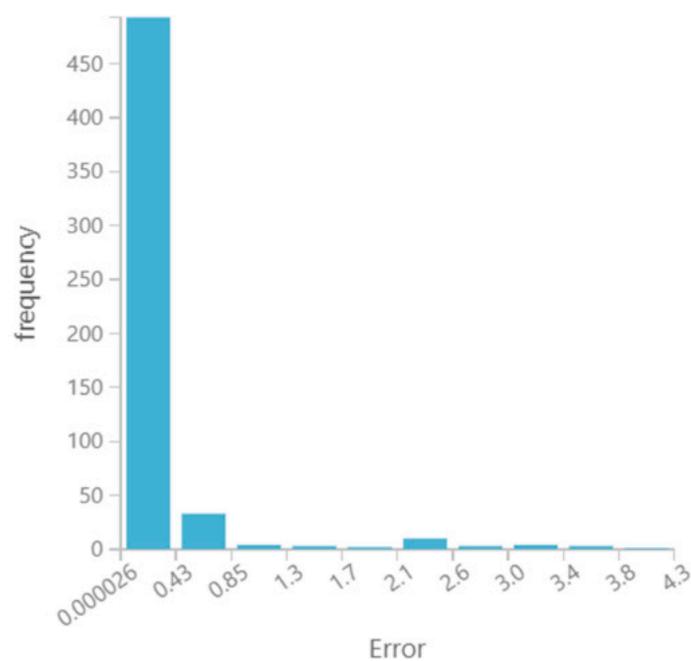
We used a dataset that was acquired from an industrial reactor to evaluate the performance of our trained model. Figure 8 shows the comparison between the predicted values and the actual values of the  $dc/dt$ . Again, the authors would like to emphasize if the predicted value of the target as a function of the “true” values of the test samples follows the “ $y = x$ ” line, one can conclude that the model performance is at its best. It is worth mentioning that we only used two features for this prediction: namely, total  $O_2$  flow and lance height. It is always one of the goals of any machine-learning development to predict the target with a fewer number of variables/predictors. This is because if the model deals with many predictor variables, then there is a high chance that there are hidden relationships between some of them, leading to redundancy and, even if there is no relationship between any of them, the model can suffer from overfitting when there are a large number of predictor variables. In addition, a model that can predict with a fewer number of predictor variables is more practical due to some considerations, such as data availability, storage, computer resources, time taken for computation, etc. Thus, this is one of the achievements of this study, that the developed model can predict the target by only using two predictor variables.

The scatter plot demonstrates that our trained model could predict the  $dc/dt$  precisely. Figure 9 shows the error histogram for this prediction. The most frequent error was 0.000026. The metrics (Table 3, column 6) were recorded to be 0.25, 0.62, 0.04 and 0.009 for the mean absolute error, root mean squared error, relative absolute error and relative squared error, respectively. The coefficient of determination was computed to be 0.98. These values confirm that our trained model can be used at industries to predict and control the variation of the  $dc/dt$  in an actual reactor. Figure 8 shows that the

machine-learning algorithm can predict the decarburization rate very accurately without employing any simplifications and without taking into account all the reactions, interactions, mass and heat transfers and fluid flows. In fact, all of these parameters are already inherited in the dataset, and an algorithm trained on the real dataset naturally learns the relationships between all of the parameters involved in the process without exactly knowing the physical meanings of them.



**Figure 8.** Scatter plot comparing the predicted values of the  $dc/dt$  using neural network method with the actual values of the  $dc/dt$  for an industrial dataset with the two operating parameters of total oxygen flow and lance height only.



**Figure 9.** Error histogram of the predicted values of the  $dc/dt$  for the industrial dataset with the two operating parameters of total oxygen flow and lance height only.

## 6. Conclusions

We applied a machine learning-based analysis to a dataset (operating and output data) from a pilot basic oxygen steelmaking (BOS) converter. Correlation analysis showed:

- A strong positive correlation between the rate of decarburization ( $dc/dt$ ) and total oxygen flow.
- A negative correlation with lance height.
- Less obviously, the decarburization also showed a positive correlation with the temperature of the waste gas,  $CO_2$  content in the waste gas and  $O_2$  in the waste gas.
- The pilot plant dataset was used for training and test data to develop a neural network-based regression to predict the decarburization rate. The developed algorithm was used successfully to predict the decarburization rate in a BOS furnace in an actual manufacturing plant based on the two operating parameters of total oxygen flow and lance height only.
- The performance was satisfactory, with a coefficient of determination of 0.98, confirming that the trained model can adequately predict the variation in the  $dc/dt$  within BOS reactors.

The method is easily scalable for industrial applications. In addition, the machine-learning model does not simplify the problem and is able to predict the decarburization rate accurately through learning from the real dataset acquired from BOS pilot plants.

**Author Contributions:** A.R. and S.S. wrote the machine-learning algorithms and produced the figures. A.R. and S.S. discussed the machine-learning results. Z.L. provided and prepared the datasets. A.R., S.S. and Z.L. discussed the materials case study. All authors wrote and commented on the manuscript and figures. All authors have read and agreed to the published version of the manuscript.

**Funding:** Z.L. would like to thank the funding from EPSRC (Engineering and Physical Sciences Research Council, UK) under grant number EP/N011368/1 (EPSRC Fellowship).

**Acknowledgments:** The authors are grateful to Johan Eriksson at Swerea MEFOS for the interpretation of the information of the trials carried out in a 6-ton pilot converter at the MEFOS in Sweden. The authors also appreciate Chris Barnes at Tata Steel, UK, who provided the actual industry BOS converter data for the validation of the work.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Ogawa, Y.; Yano, M.; Kitamura, S.; Hirata, H. Development of the continuous dephosphorization and decarburization process using BOF. *Steel Res. Int.* **2003**, *74*, 70–76. [[CrossRef](#)]
2. Millman, M.S.; Kapilashrami, A.; Bramming, M.; Malmberg, D. *IMPHOS—Improving Phosphorus Refining*; European Commission, Research Fund for Coal and Steel: Luxembourg, 2011; ISSN 1831-9424.
3. Jung, I.H.; Hudon, P.; Van Ende, M.A.; Kim, W.Y. Thermodynamic database for  $P_2O_5$ -containing slags and its application to the dephosphorization process. In Proceedings of the Iron & Steel Technology Conference, Indianapolis, IN, USA, 5–8 May 2014; Volume 1, pp. 1257–1268.
4. Rout, B.; Brooks, G.; Rhamdhani, M.A.; Li, Z.; Schrama, F.N.; Sun, J. Dynamic model of basic oxygen steelmaking process based on multi-zone reaction kinetics: Model derivation and validation. *Metall. Mater. Trans. B* **2018**, *49*, 537–557. [[CrossRef](#)]
5. Ersson, M.; Hoglund, L.; Tilliander, A. Dynamic coupling of computational fluid dynamics and thermodynamics software: Applied on a top blown converter. *ISIJ Int.* **2008**, *48*, 147–153. [[CrossRef](#)]
6. Cox, I.J.; Lewis, R.W.; Ransing, R.S. Application of neural computing in basic oxygen steelmaking. *J. Mater. Process. Technol.* **2002**, *120*, 310–315. [[CrossRef](#)]
7. Das, A.; Maiti, J.; Banerjee, R.N. Process control strategies for a steelmaking furnace using ann with Bayesian regularization and anfis. *Expert Syst. Appl.* **2010**, *37*, 1075–1085. [[CrossRef](#)]
8. Han, M.; Zhao, Y. Dynamic control model of BOF steelmaking process based on anfis and robust relevance vector machine. *Expert Syst. Appl.* **2011**, *38*, 14786–14798. [[CrossRef](#)]
9. Kubat, C.; Takin, H.; Artir, R.; Yilmaz, A. Bofy-fuzzy logic control for the basic oxygen furnace. *Robot. Auton. Syst.* **2004**, *49*, 193–205. [[CrossRef](#)]
10. Wang, X.; Han, M.; Wang, J. Applying input variables selection technique on input weighted support vector machine modeling for BOF end point prediction. *Eng. Appl. Artif. Intell.* **2010**, *23*, 1012–1028. [[CrossRef](#)]

11. Xu, L. A model of basic oxygen furnace (BOF) endpoint prediction based on spectrum information of the furnace flame with support vector machine (SVM). *Optik* **2011**, *122*, 594–598. [CrossRef]
12. Wang, X.; Xing, J.; Dong, J.; Wang, Z. Data driven based endpoint carbon content real time prediction for BOF steelmaking. In Proceedings of the 36th Chinese Control Conference (CCC), Dalian, China, 26–28 July 2017; pp. 9708–9713.
13. Liu, Y.; Zhao, T.; Ju, W.; Shi, S. Materials discovery and design using machine learning. *J. Mater.* **2017**, *3*, 159–177. [CrossRef]
14. Ward, L.; O’Keeffe, S.C.; Stevick, J.; Jelbert, G.R.; Aykol, M.; Wolverton, C. A machine learning approach for engineering bulk metallic glass alloys. *Acta Mater.* **2018**, *159*, 102–111. [CrossRef]
15. Li, W.; Jacobs, R.; Morgan, D. Predicting the thermodynamic stability of perovskite oxides using machine learning models. *Comput. Mater. Sci.* **2018**, *150*, 454–463. [CrossRef]
16. Cassar, D.R.; de Carvalho, A.C.; Zanutto, E.D. Predicting glass transition temperatures using neural networks. *Acta Mater.* **2018**, *159*, 249–256. [CrossRef]
17. Meredig, B.; Agrawal, A.; Kirklin, S.; Saal, J.E.; Doak, J.W.; Thompson, A.; Zhang, K.; Choudhary, A.; Wolverton, C. Combinatorial screening for new materials in unconstrained composition space with machine learning. *Phys. Rev. B* **2014**, *89*, 094104. [CrossRef]
18. Rahnama, A.; Clark, S.; Sridhar, S. Machine learning for predicting occurrence of interphase precipitation in HSLA steels. *Comput. Mater. Sci.* **2018**, *154*, 169–177. [CrossRef]
19. Ramprasad, R.; Batra, R.; Piliand, G.; Mannodi-Kanakithodi, A.; Kim, C. Machine learning in materials informatics: Recent applications and prospects. *Npj Comput. Mater.* **2017**, *3*, 1–13. [CrossRef]
20. Available online: <https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/multiclass-neural-network> (accessed on 2 January 2020).
21. Chattopadhyay, K.; Isac, M.; Guthrie, R.I.L. Applications of Computational Fluid Dynamics (CFD) in iron- and steelmaking (part I & part II). *Ironmak. Steelmak.* **2010**, *37*, 554–569.
22. Ek, M.; Shu, Q.F.; van Boggelen, J.; Sichen, D. New approach towards dynamic modelling of dephosphorisation in converter process. *Ironmak. Steelmak.* **2012**, *39*, 77–84. [CrossRef]
23. Oguchi, S.; Robertson, D.G.C.; Deo, B.; Grieveson, P.; Jeffes, J.H.E. Simultaneous dephosphorization and desulphurization of molten pig iron. *Ironmak. Steelmak.* **1984**, *11*, 202–213.
24. Spooner, S.; Rahnama, A.; Warnett, J.M.; Williams, M.A.; Li, Z.; Sridhar, S. Quantifying the Pathway and Predicting Spontaneous Emulsification during Material Exchange in a Two Phase Liquid System. *Sci. Rep.* **2017**, *7*, 14384. [CrossRef]
25. Spooner, S.; Li, Z.; Sridhar, S. Spontaneous emulsification as a function of material exchange. *Sci. Rep.* **2017**, *7*, 5450. [CrossRef] [PubMed]
26. Dogan, N.; Brooks, G.A.; Rhamdhani, M.A. Kinetics of decarburization reaction in oxygen steelmaking process. In *High Temperature Processing Symposium*; Swinburne University of Technology: Melbourne, Australia, 2010; pp. 9–11.
27. The AISE Steel Foundation. *The Making, Shaping and Treating of Steel (Steelmaking and Refining Volume)*; AIST: Pittsburgh, PA, USA, 1998; pp. 475–524.

