Received: 3 May 2024 | Revised: 15 July 2024 | Accepted: 20 August 2024 | Published online: 27 August 2024

RESEARCH ARTICLE

Archives of Advanced Engineering Science
2024, Vol. 00(00) 1–8
DOI: 10.47852/bonviewAAES42023358

BON VIEW PUBLISHING

Comparison Between Empirical Strategies for Predicting Endpoint Phosphorus Content in BOF Steelmaking Process

Diego Henrique de Souza Chaves[1], Iara Campolina Dias Duarte[2], Esly Ferreira da Costa Junior[2,3,*] and Andréa Oliveira Souza da Costa[2,3]

[1]*Control and Automation Area, Federal Institute of Minas Gerais, Brazil*

[2]*Graduate Program in Mechanical Engineering, Federal University of Minas Gerais, Brazil*

[3]*Graduate Program in Chemical Engineering, Federal University of Minas Gerais, Brazil*

Abstract: Dephosphorization is a reaction of important role in steelmaking process, and the correct adequacy of endpoint phosphorus content would improve the quality and productivity of steel in basic oxygen furnace (BOF) processing. Aiming to meet the required steel specifications and reduce process time, two different empirical strategies were established for predicting the endpoint phosphorus content in BOF steelmaking process: linear regression and neural network. Eight variables that affect the endpoint phosphorus content (selected as output) were determined as the input variables of the models. The performances of predictions were evaluated simultaneously with the sensitivity analysis of the model to variations in the values of its input variables. Sensitivity analysis is essential as it reveals the impact of input variables on results, although it is often neglected due to its complexity and the need for multiple simulations. Integrating sensitivity analysis with prediction techniques allows for identifying key variables and making decisions. Both empirical models are suitable and reliable for decision-making in the process and can be used as tools for predicting the endpoint phosphorus content, where the neural network has higher accuracy. The sensitivity analysis showed that the two variables that most affect the response of the empirical models were the percentage of oxygen volume of oxygen blown until the sub-lance in relation to the estimated total volume, and the phosphorus concentration in the sub-lance.

Keywords: dephosphorization, linear regression, neural network, sensitivity analysis

1. Introduction

Steelmaking in basic oxygen furnace (BOF) is a complex physical-chemical process where molten iron from a blast furnace, along with scrap steel, is refined into high-quality steel [1, 2]. During the process, components such as carbon, phosphorus, sulfur, and manganese are oxidized while the metal bath temperature is raised. This oxidation occurs due to the injection of oxygen through a vertical lance inserted into the converter [2, 3].

The BOF converter has advantageous thermodynamic and kinetic conditions in conducting these reactions in an environment of high basicity [1, 2, 4–6].

The main process indicators, such as chemical compositions, and temperature of the metal bath are measured and analyzed during and after the refinement step, and they determine if the quality of steel is in accordance with the desired specification [1–3]. Phosphorous removal is an exothermic reaction of extreme importance in the steelmaking process. Equation (1) represents the exothermic dephosphorization chemical reaction [6] where parentheses and brackets denote components in the slag and metal, respectively.

$$2[P] + 5(FeO) + 4(CaO) \leftrightarrow (4CaO \cdot P_2O_5) + 5[Fe] \qquad (1)$$

Dephosphorization occurs at the metal-slag interface where phosphorus is oxidized by reacting with iron oxide in the slag. The formed phosphorus pentoxide then associates with lime and moves into the slag. The rate of dephosphorization increases with the increase in CaO, thus allowing the stability of P_2O_5 in highly basic slags saturated with CaO [7].

Phosphorus is considered an impurity, and its content must comply with strict specifications due to harmful effects on the mechanical properties of steel [3–6, 8]. Controlling the steel's composition is essential to ensure the material meets the required standards of strength, ductility, and durability.

Currently, there are mainly two methods to control the endpoint phosphorus content in BOF [7]. One of the methods is the laboratory analysis of samples collected directly from the metal bath by inserting a sub-lance into the furnace, sampling occurs during and at the end of the batch process; however, the results of these analysis are time-consuming, and then, the production efficiency

*__Corresponding author:__ Esly Ferreira da Costa Junior, Graduate Program in Mechanical Engineering, Federal University of Minas Gerais, Brazil and Graduate Program in Chemical Engineering, Federal University of Minas Gerais, Brazil. Email: esly@deq.ufmg.br

is compromised. The other method is to predict phosphorus composition using predictive models.

Although there are works that compare empirical models obtained by linear regression and neural networks [7], they do not develop the sensitivity analysis of the obtained models to variations in the inputs for this process. Thus, the objective of this study is to conduct a sensitivity analysis on two empirical strategies for predicting the endpoint phosphorus content in BOF: Regression and neural network. This investigation enables determining how variations in conditions affect the quality of predictions.

2. Literature Survey

The use of empirical models and data-driven approaches is widely employed in industrial applications [9–11], and this scenario is no different in the steel industry [12–15]. This is primarily due to two factors. One is the complexity of physical-chemical phenomena, which limits the use of purely mathematical models, and the other is the availability of large amounts of process data, because of technological advancements in data management systems, informatics, control, and instrumentation.

The role of mathematical models is especially important to guide the operation of steel production, but they have some inherent limitations in practical applications because BOF steelmaking process has simultaneously chemical reactions, heat transfer, and mass transfer. The complexity of BOF steelmaking entails its modeling and control difficult to be reduced to sets of equations [2, 6, 7, 14, 16].

In addition to the use of an automatic sub-lance, advances in data acquisition systems, computational performance, and the development of measurement techniques, sensors, and devices have favored the availability of information that can be used to propose intelligent controllers and dynamic non-linear solutions to improve the monitoring of the process [2].

Accurately predicting the phosphorus content of steel at the endpoint of the BOF is essential to increase productivity and reduce production costs since dephosphorization cannot occur in subsequent steps in BOF steelmaking process [7, 17]. Any deviation between the phosphorus content and the desired specification causes reprocessing and corrective actions, increasing the production time and, therefore, should be avoided. This justifies the several research that have been done in this field of study [13, 15, 18–20].

An empirical model of endpoint steel phosphorus content in BOF was proposed [7] based on the principal component analysis and backpropagation (BP) neural network. The combined model achieved significant results, more than 86% agreement between the data predicted by the model and the experimental data. BP neural network was also used for controlling endpoint of steel phosphorus content in the BOF converter through theoretical analyses of the dephosphorization process [18]. The Levenberg-Marquardt (LM) method as a training algorithm can also result in satisfactory results for endpoint phosphorus prediction [10, 21].

Other strategies can be combined to obtain empirical models. Thermodynamic analysis of the steel dephosphorization reaction can be used to develop an empirical model based on multivariate regression analysis to predict the endpoint phosphorus content [6]. A predictive model was proposed [16] based on computer vision and general regression neural network, using flame image processing and pattern classification technologies. An optimization algorithm may be used for training a neural network with the aim of minimizing the endpoint phosphorous content and maximizing the bath temperature [5].

3. Material and Methods

Creating a highly accurate empirical model relies on having a comprehensive dataset that accurately reflects the system being studied. The careful selection of input and output variables significantly influences the predictive capability and overall performance of the model [7, 13, 15, 18–20]. The 8 input and the output variables are presented in Table 1 with the respective average dimensionless values.

A survey was conducted on all operational variables available in the process based on metallurgical theory and research published [7, 18, 20, 22]. To obtain the empirical model for the BOF endpoint phosphorus content (output variable), the choice of the input variables was determined by its influence on the phosphorus content. After this survey, the number of input parameters has been reduced to eight.

The processed dataset was obtained through a query to the SQLServer database of a Brazilian steel company. In the studied steel production process, some information such as temperature, carbon content, and blown oxygen volume are automatically acquired by the supervisory system. On the other hand, other information like phosphorus content, manganese, sulfur, and other chemical compositions is manually entered into the system by the operational team. The analysis of samples from the metal bath is conducted by the responsible laboratory, and these data are not automatically available in the supervisory system.

The database (4902 runs) was collected in a Brazilian steelmaking industry and runs that had reprocessing, for whatever reason, were excluded from the database. The nondimensionalization was

Table 1
Dataset of operating variables of the case study. Normalized values [0, 1]

Code	Symbol	Description	Mean	Standard deviation
C_SL	X1	Carbon concentration at sub-lance	0.349	0.190
P_TARGET	X2	Target phosphorus concentration	0.557	0.232
VOL_SL	X3	Percentage of O_2 volume blown until the sub-lance in relation to the estimated total volume VOL_EST	0.713	0.091
VOL_END	X4	Percentage of O_2 volume estimated to blow after sub-lance	0.357	0.123
T_END	X5	Endpoint temperature	0.492	0.201
D_CAL	X6	Calcined dolomite weight	0.289	0.127
DOL	X7	Raw dolomite weight	0.433	0.150
P_SL	X8	Phosphorus concentration at sub-lance	0.425	0.186
P_END	Y	Endpoint phosphorus concentration	0.426	0.151

Table 2
Variable ranges and units

Code	Minimum	Maximum	Unit
C_SL	0.032	0.999	ppm
P_TARGET	0.010	0.040	ppm
VOL_SL	0.4810	0.9957	%
VOL_END	0.002	0.019	%
T_END	1610	1734	ºC
D_CAL	1024	13749	ton
DOL	1028	21019	ton
P_SL	0.010	0.083	ppm
P_END	0.001	0.039	ppm

performed using Equation (2) and relates each variable to its maximum and minimum values, resulting in dimensionless values to facilitate the comparison and analysis of phenomena.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2}$$

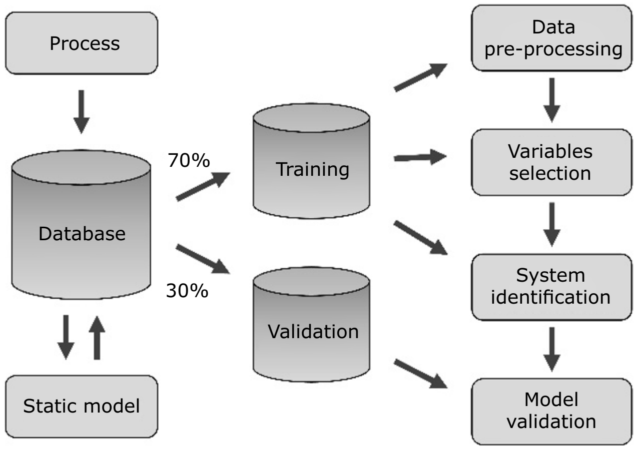
Table 2 presents the variables along with their respective units, maximum, and minimum values.

Figure 1 illustrates the process of obtaining the experimental values of the variables presented in Tables 1 and 2. The figures in this article were created using MS Office and MATLAB. The steel specification is used to run the static models in the operation room, which estimate the volume of O_2 to be injected into the furnace (VOL_EST). Then, the blowing of oxygen starts, and fluxing materials are added to assist the steel oxidation. Temperature, O_2 volume, and carbon concentration are real-time and online measurements, while phosphorus concentrations (at sub-lance and endpoint) are analyses carried out in the laboratory, taking a few minutes for the results to be available.

The model adjusts its parameters to minimize the quadratic error between predictions and actual results in the training data. Validation data are used to evaluate the performance of the model on a dataset independent of the training data to avoid overfitting. The two strategies analyzed in this work followed the following methodology represented in Figure 2.

The experimental database was randomly subdivided into two: training, corresponding to 70% of the data, and validation, the remaining 30%. Two phosphorus empirical prediction strategies were analyzed: (1) multiple linear regression (MLR) model and (2) neural network. The analysis of the performance of the models was based on the validation data. Sensitivity analysis for both strategies was performed to understand the effect of variation that each input produces on phosphorus prediction. Through this analysis, based on a methodology similar to Cvetković et al. [23],

Figure 2
Methodology steps for obtaining the empirical models



it is possible to identify critical variables and steps for proposing improvements in the BOF steelmaking process.

3.1. Neural network

The Neural Network Toolbox was used, which is the primary toolbox for working with neural networks in MATLAB. The BP is the most widely used algorithm in ANN (artificial neural network), which is a gradient descent method designed to minimize the total error (or mean error) of the output computed by the network [22]. Traditional BP algorithm model has some issues with numerical convergence speed and, the objective function can be easily stick in the local minima [21, 22]. The LM algorithm is suggested for overcoming these difficulties. LM algorithm is an improved modality of Gauss-Newton method, and it has advantages, such as global property of gradient method and faster convergent speed for medium-size neural networks [21].

3.2. Linear regression

The MLR adopts more than one explanatory (independent) variable to predict the response (dependent) variable. This is useful when you want to understand how multiple factors simultaneously affect the dependent variable. The polynomial regression is used to describe the relationship between the inputs and the output variables as an n^{th} degree polynomial [9, 24]. MLR is a relevant tool to propose experimental routines of complex systems [25], such as BOF steelmaking process since these models do not require a deep knowledge of the involved phenomena in the processes.

Figure 1
Methodology chart for the experimental data acquisition. *Laboratory analysis

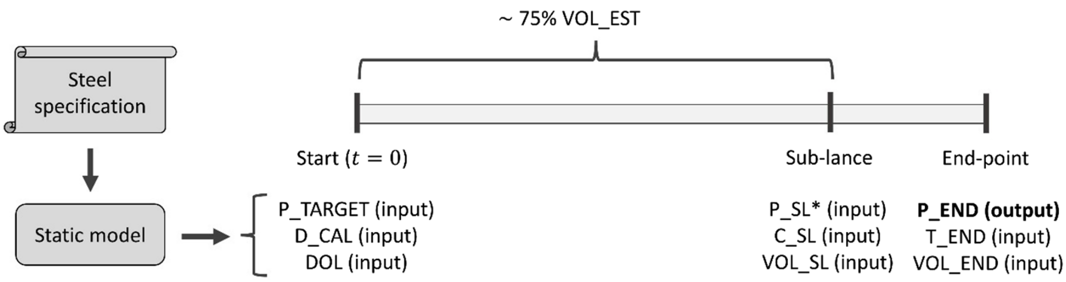


Table 3
Correlation analysis of the process variables for regression data

	X1	X2	X3	X4	X5	X6	X7	X8	Y
X1	1.00	–	–	–	–	–	–	–	–
X2	0.10	1.00	–	–	–	–	–	–	–
X3	−0.59	0.06	1.00	–	–	–	–	–	–
X4	0.58	−0.06	–	1.00	–	–	–	–	–
X5	0.08	0.33	−0.23	0.27	1.00	–	–	–	–
X6	−0.03[a]	−0.61	−0.10	0.13	0.05	1.00	–	–	–
X7	0.01[a]	−0.23	−0.12	0.14	0.15	0.32	1.00	–	–
X8	0.37	0.47	0.01[a]	−0.01[a]	0.19	−0.33	−0.04	1.00	–
Y	0.19	0.66	0.16	−0.16	0.36	−0.48	−0.15	0.73	1.00

[a]Not significant correlations (*p*-value > 0.05)

4. Results and Discussion

To perform linear regression and neural network training, it is necessary to assume that variations in the input variables affect the output variable, and the variables selected for the problem were evaluated using the correlation matrix (Table 3). The obtained values from this analysis reflect the degree of association between two or more variables of interest. A value close to 1 indicates a strong positive correlation, and a value close to −1 indicates also a strong, negative correlation, though. A value close to zero indicates a weak or non-existent correlation represented by cases in which the calculated p-values are greater than 5% (test significance) [12, 24].

All input variables are significantly correlated (*p*-value < 5%) to the output variable (endpoint phosphorous content). The three input variables that most correlate with the output are target phosphorus concentration, phosphorus concentration at sub-lance, and calcined dolomite weight. A direct and strong correlation is observed between all phosphorus concentrations present in the model. The strong relationship of calcined dolomite is justified by the need to adjust the basicity of the metallic bath to optimize the dephosphorization [7, 8].

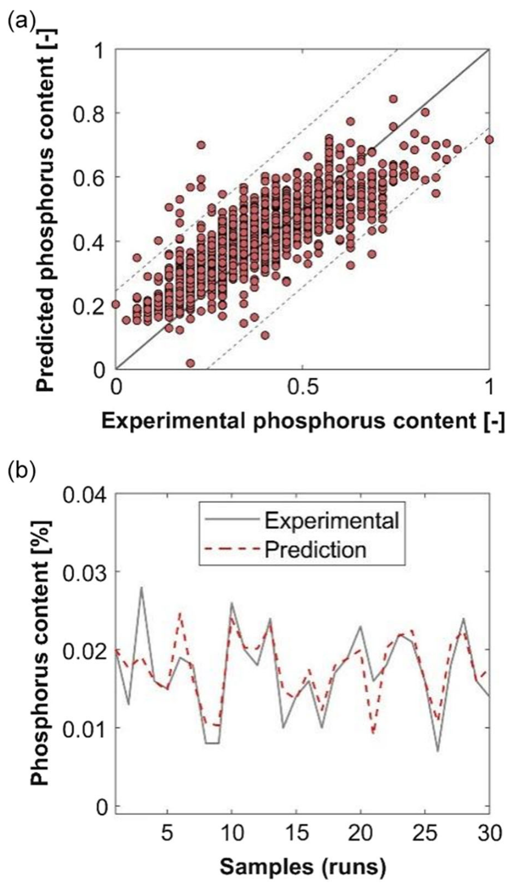
The phosphorus content in the metal bath varies during the process. Initially, the dephosphorization reactions occur due to the combined effects of oxygen injection and the addition of fluxing materials. These materials are important to adjust the basicity of slag to maintain the equilibrium condition of residual phosphorus. However, oxidation reactions favor raising the temperature, and this destabilizes the equilibrium conditions of dephosphorization reactions, regenerating phosphorus to steel [8]. To avoid this problem, the operation must control the endpoint temperature and endpoint phosphorus content of the metallic bath within the specified values for the type of steel being produced.

4.1. Neural network

For the configuration of the neural network structure, 1 to 5 intermediate layers were tested, with up to 20 neurons in each. Two training algorithms for the neural network were evaluated: BP and LM. The network was trained using the training data based on the performance mean squared error (MSE) which is directly related to the R^2, the smaller the MSE, the higher the R^2. The best configuration obtained was: 3 intermediate layers of 5, 5, and 10 neurons, respectively, using the LM algorithm with 100 epochs of training.

Figure 3a shows the graph of the values predicted by the neural network in relation to the experimental values of phosphorus concentration at the end of the process using the validation data.

Figure 3
Model validation: (a) Parity relation between the predicted values by the neural model and the factory-collected variable (experimental) values; (b) trend plots for the experimental values and the neural model predictions



Some discrepancies between predicted and actual data can be identified, and it is possible to analyze the quality of predictions.

The neural network model had an R^2 for validation of 69.5%, indicating acceptable agreement between the model and the data used. In industrial process, there are several types of steel being produced in the BOF converter. The data were obtained without selecting any specific type of steel and therefore, this causes the data dispersion to become slightly larger. This predictive model can even reduce this data uncertainty when implemented in the process. Zhou et al. [18] also achieved good results using BP with ANNs when developing their prediction model for controlling the phosphorus content.

Figure 3b compares the neural model prediction to the experimental phosphorus concentration. Neural networks have the characteristic of being non-linear [2, 7]. A better fit (higher R^2) of the neural network model than the regression model could be justified on the basis that the neural network could adapt to the non-linearities of the dephosphorization process.

It is important to emphasize that the predicted endpoint phosphorus values, in none of the events, exceeded the target endpoint phosphorus value specified for the steel, even in cases where the results predicted by the neural model were greater than the experimental endpoint phosphorus values. The proposed model is reliable for decision-making since the limit of acceptable phosphorus content for that steel is being respected. Phosphorus removal is impossible in subsequent processes. If the phosphorus content exceeds the control standard, that steel will be reused as scrap.

Figure 4a shows how the neural network responds to percentage changes in the input variables most correlated with the output (See Table 3). It is possible to see a positive relationship for the sensitivity analysis in the variables X3 (VOL_SL), X8 (P_SL), X2 (P_TARGET), and X5 (T_END) and a negative one in X6 (D_CAL), which agrees with the correlation analysis presented in Table 3.

Sensitivity analysis was evaluated as the percentage change in the model response variable from −20% to +20% changes in the mean values of each of the input variables, keeping the others with mean constant. The ±20% value was strategically chosen according to the average standard deviation of the process variables.

The maximum variation of the output variable (endpoint phosphorus content) was −3% and +4% for a maximum variation, respectively, of −20% and +20% in variable X3 (VOL_SL), showing a non-symmetric sensitivity profile. The asymmetric sensitivity profile can also be observed in Figure 4b, which represents a Tornado chat which facilitates the graphical analysis of the sensitivity, but it does not allow the visualization of non-linearities in the sensitivity. In addition to the asymmetrical profile, Figure 4a shows a non-linear profile for X3 (VOL_SL) and X2 (P_TARGET). Some variables such as X5 (T_END), X6 (D_CAL), and X8 (P_SL) present a percentage variation profile closer to linearity, that is, closer to a straight line.

4.2. Linear regression

The data used for regression/training and those used for validation are the same for the two empirical models used (neural network and regression) to make it possible the comparison between the two models' predictions and sensibility analyses. For the regression model, Figure 5a shows a greater dispersion

Figure 4
Neural model sensitivity analysis: (a) Graph of variations in phosphorus content due to variations in the inputs between −20% to +20% of their base value (95% confidence interval); (b) Tornado diagram for −20% and +20% variations

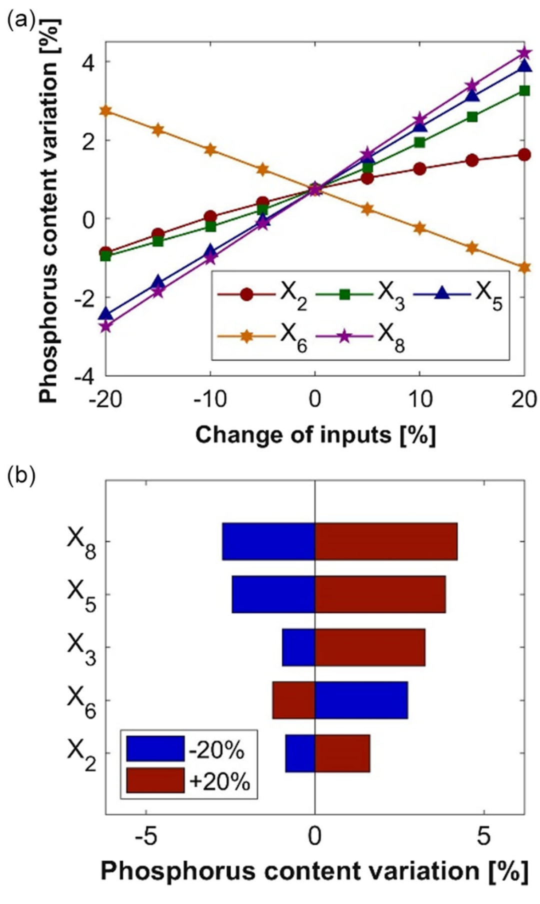


Figure 5
Model validation: (a) Parity relation between the predicted values by the regression model and the factory-collected variable (experimental) values (95% confidence interval); (b) trend plots for the experimental values and the regression model predictions

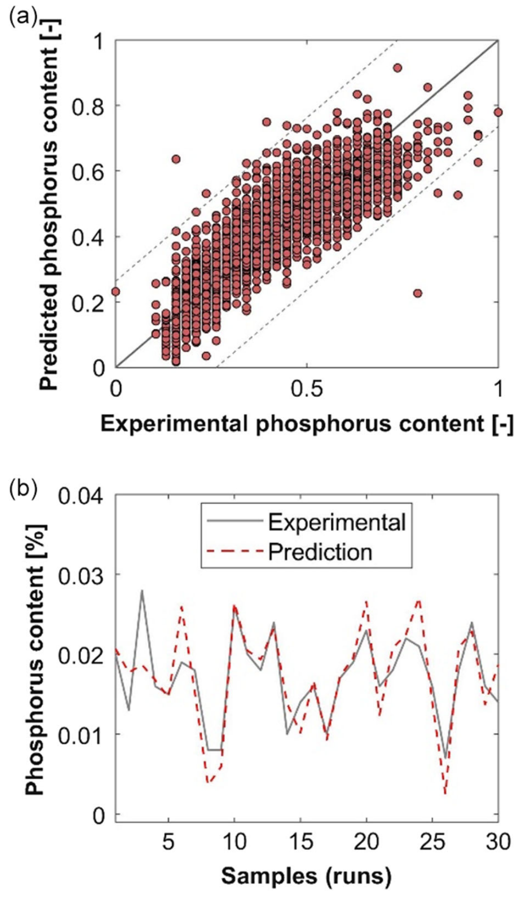
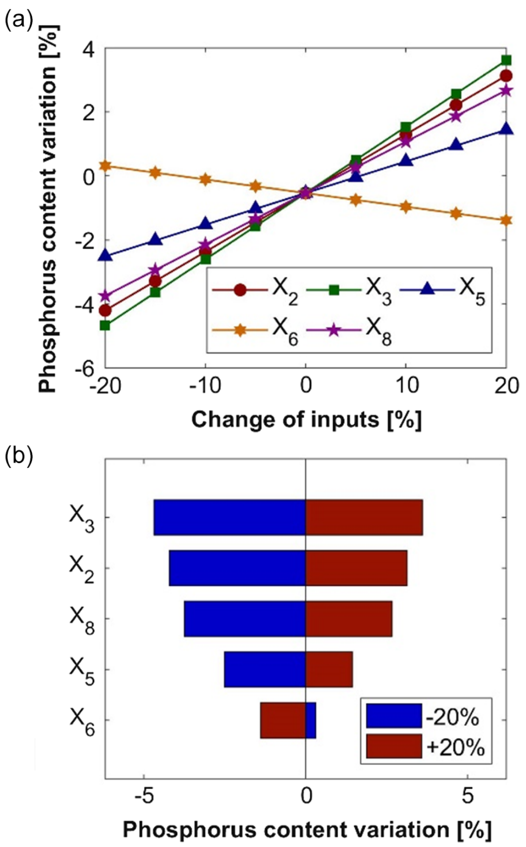


Figure 6
Regression model sensitivity analysis: (a) Graph of variations in phosphorus content due to variations in the inputs between −20% to +20% of their base value; (b) Tornado diagram for −20% and +20% variations



between the predicted and observed data compared with the neural model (compare Figures 3 and 5), impacting the R^2 value, which was now 64.7%.

Despite a lower R^2 coefficient (higher MSE) for the regression model, it is important to point out that no predicted value was greater than the target endpoint phosphorus value specified for the steel. Regarding only the coefficient R^2, both empirical models, regression model and neural network, are suitable for decision-making in the process and can be used as tools for predicting the endpoint phosphorus content aiming for time and cost savings, where the neural network has higher accuracy.

The analysis of the parity graph revealed that the performance of the neural network model was significantly superior compared to the linear regression approach. The results highlight the neural network's ability to capture non-linear patterns that other techniques usually do not capture. When analyzing the first 30 runs (Figure 5b), it is evident that the performance of the regression model was inferior compared to the neural network model (Figure 3b). While the graph may be useful for identifying trends, it did not fit the data as well as the neural network model.

Sensitivity analysis for the regression model shows a more linear sensitivity profile (Figure 6a) in addition to a more symmetrical profile (Figure 6b), both compared to the neural network model sensitivity.

Table 4
Prediction analysis for the endpoint phosphorus: Factory-collected values against the estimates given by the models using the validation dataset. *Normalized values [0, 1]

Descriptive statistics*	Factory-collected variable (y)	Neural model	Regression model
Fitted values ($\hat{y}$)			
Mean	0.4258	0.3647	0.3593
Standard deviation	0.1508	0.1393	0.1440
R^2 validation	–	0.6953	0.6471
Residuals			
Mean	–	0.0541	0.0008
Standard deviation	–	0.0755	0.0939

As widely discussed, neural networks have a highly non-linear characteristic in their response variables [7], which directly interferes with the model's sensitivity to variations in its inputs.

A linear regression model has a characteristic of linearity in its parameters, which may have impacted the profiles shown in Figure 6. This linear and symmetric sensitivity is desirable in decision-making processes because of the characteristic of unbiased representation. Symmetry also allows for easier maintenance and simpler control techniques. Table 4 summarizes the performances of the two models obtained. Residuals for both empirical models were analyzed and have a null mean and are independent.

From the results summarized in Table 4, while the neural network model has better accuracy in predicting phosphorus content endpoint, a fact also pointed out by Subramanyam and Narayanan et al. [10], the regression model has approximately symmetric and linear sensitivity. Other databases for the same process need to be studied to analyze and validate the sensitivity analysis, but for the data used, the neural model was better in an overall aspect.

For linear regression strategy, the significance of the regressed parameters was verified, and those that were not significant were removed (p-value < 0.05). This is a common practice in statistical analysis to ensure that only meaningful variables are included in the final model. The results obtained by Zhou et al. [18] corroborate the findings in Table 4. These authors also achieved good fits using these empirical strategies, with R-squared coefficients of 0.83 and 0.81 for the neural network and regression, respectively. The higher R-squared value for the neural network indicates superior performance in prediction compared to regression.

Mahanta et al. [15] also developed models describing process parameters such as temperature and carbon and phosphorus contents at the end of the blow, correlating influential variables with phosphorus content at the endpoint. Their results align with the sensitivity analyses conducted by the authors.

Improvement strategies can be implemented for both empirical models. For example, combination of systems such as evolutionary neural network, bi-objective genetic programming, and evolutionary deep neural network generated models that described three essential process parameters, temperature, carbon, and phosphorus contents of the metal bath in BOF steelmaking process [15]. Regression models can be improved using combinations of multiple variables and their inverses to increase the predictive capacity of simple linear models [25]. Furthermore, higher-order polynomials can yield good results, being able to highlight the good performance of polynomials of the second order [9, 24] and fourth order [25].

5. Conclusion

Dephosphorization is a reversible exothermic reaction present in the steelmaking process and has non-linear characteristics, making the process complex. Thus, since the relationships between variables are not fully understood, empirical models can capture hidden patterns and trends in the data, allowing a better understanding of the process. The two empirical models studied, the regression model and the neural network, were adequate to predict the endpoint phosphorus content in the BOF steelmaking process. This variable is important, given that laboratory analyses require considerable process time and it is not possible to adjust this concentration in subsequent steps.

The proposed linear regression model was compared with the neural network in predicting phosphorus content and conducting sensitivity analysis. Furthermore, the regression linear model is unique, including variables not found in the literature, such as VOL_SL and VOL_END. Despite the superior performance of the neural network, a fact due to its non-linear response characteristic, this model better predicted the response variable ($R^2 = 69.5\%$). However, the sensitivity analysis showed an asymmetric and non-linear behavior. The methodology used to develop the proposed model by the authors can be applied to any other mineral processing method. However, the two predictive phosphorus models are only applicable to steelmaking processes that utilize the BOF converter.

Enhancements can be implemented to improve accuracy, which is intrinsically related to instrumentation accuracy and process decision-making. The results suggest that possible improvements in the process should consider a better precision in the activation of sub-lance, given that the percentage of oxygen volume of oxygen blown until the sub-lance in relation to the estimated total volume has been shown to have a greater influence on the sensitivity of the model's output variable.

Funding Support

This work was partially funded by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), finance code 001, and by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brazil, 312248/2022-9).

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data available on request from the corresponding author upon reasonable request.

Author Contribution Statement

Diego Henrique de Souza Chaves: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization; **Iara Campolina Dias Duarte:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization; **Esly Ferreira da Costa Junior:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition; **Andréa Oliveira Souza da Costa:** Conceptualization, Investigation, Resources, Data curation, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

List of Abbreviations

ANN	Artificial Neural Network
BOF	Basic Oxygen Furnace
BP	Backpropagation
LM	Levenberg-Marquardt
MLR	Multiple Linear Regression
MSE	Mean Squared Error

References

[1] Ferreira, A. F., Ferreira, L. O., & Alcântara Junior, Z. (2014). Predictions at the blow end of the LD-KGC converter by a semi-dynamic control model. *International Journal of Recent Advances in Mechanical Engineering*, *3*(1), 17–31.

[2] Han, M., & Zhao, Y. (2011). Dynamic control model of BOF steelmaking process based on ANFIS and robust relevance vector machine. *Expert Systems with Applications*, *38*(12), 14786–14798. https://doi.org/10.1016/j.eswa.2011.05.071

[3] Tomiyama, S., Uchida, Y., Mizuno, H., Akiu, K., & Maeda, T. (2015). A novel control algorithm for dephosphorization in an LD converter. *Journal of Process Control*, *25*, 35–40. https://doi.org/10.1016/j.jprocont.2014.11.002

[4] Heo, J. H., & Park, J. H. (2018). Effect of direct reduced iron (DRI) on dephosphorization of molten steel by electric arc furnace slag. *Metallurgical and Materials Transactions B*, *49*(6), 3381–3389. https://doi.org/10.1007/s11663-018-1406-5

[5] Pal, S., & Halder, C. (2016). Optimization of phosphorous in steel produced by basic oxygen steel making process using multi-objective evolutionary and genetic algorithms. *Steel Research International*, *88*(3), 1600193. https://doi.org/10.1002/srin.201600193

[6] Wang, Z., Xie, F., Wang, B., Liu, Q., Lu, X., Hu, L., & Cai, F. (2014). The control and prediction of end-point phosphorus content during BOF steelmaking process. *Steel Research International*, *85*(4), 599–606. https://doi.org/10.1002/srin.201300194

[7] He, F., & Zhang, L. (2018). Prediction model of end-point phosphorus content in BOF steelmaking process based on PCA and BP neural network. *Journal of Process Control*, *66*, 51–58. https://doi.org/10.1016/j.jprocont.2018.03.005

[8] Drain, P. B., Monaghan, B. J., Longbottom, R. J., Chapman, M. W., Zhang, G., & Chew, S. J. (2018). Phosphorus partition and phosphate capacity of basic oxygen steelmaking slags. *ISIJ International*, *58*(11), 1965–1971. https://doi.org/10.2355/isijinternational.isijint-2018-129

[9] Ajona, M., Vasanthi, P., & Vijayan, D. S. (2022). Application of multiple linear and polynomial regression in the sustainable biodegradation process of crude oil. *Sustainable Energy Technologies and Assessments*, *54*, 102797. https://doi.org/10.1016/j.seta.2022.102797

[10] Subramanyam, R., & Narayanan, M. (2023). Artificial neural network modeling for drying kinetics of paddy using a cabinet tray dryer. *Chemical Industry and Chemical Engineering Quarterly*, *29*(2), 87–98. https://doi.org/10.2298/ciceq220106017s

[11] Corrêa, L. Q., Bagnis, D., Franco, P. R. M., Costa Jr, E. F., & Oliveira, A. S. C. (2023). Evaluating energy generation of a building-integrated organic photovoltaic vertical façade: A case study of Latin America's pioneering installation. *SSRN*. https://doi.org/10.2139/ssrn.4546944

[12] Duarte, I. C. D., Almeida, G. M. de, & Cardoso, M. (2020). Heat-loss cycle prediction in steelmaking plants through artificial neural network. *Journal of the Operational Research Society*, *73*(2), 326–337. https://doi.org/10.1080/01605682.2020.1824552

[13] Chang, S., Zhao, C., Li, Y., Zhou, M., Fu, C., & Qiao, H. (2021). Multi-channel graph convolutional network based end-point element composition prediction of converter steelmaking. *IFAC-PapersOnLine*, *54*(3), 152–157. https://doi.org/10.1016/j.ifacol.2021.08.234

[14] Gu, M., Xu, A., Yuan, F., He, X., & Cui, Z. (2021). An improved CBR model using time-series data for predicting the end-point of a converter. *ISIJ International*, *61*(10), 2564–2570. https://doi.org/10.2355/isijinternational.isijint-2020-687

[15] Mahanta, B. K., Gupta, P., Mohanty, I., Roy, T. K., & Chakraborti, N. (2023). Evolutionary data driven modeling and tri-objective optimization for noisy BOF steel making data. *Digital Chemical Engineering*, *7*, 100094. https://doi.org/10.1016/j.dche.2023.100094

[16] Liu, H., Wang, B., & Xiong, X. (2014). Basic oxygen furnace steelmaking end-point prediction based on computer vision and general regression neural network. *Optik*, *125*(18), 5241–5248. https://doi.org/10.1016/j.ijleo.2014.05.004

[17] Qi, L., Liu, H., Xiong, Q., & Chen, Z. (2021). Just-in-time-learning based prediction model of BOF endpoint carbon content and temperature via vMF mixture model and weighted extreme learning machine. *Computers & Chemical Engineering*, *154*, 107488. https://doi.org/10.1016/j.compchemeng.2021.107488

[18] Zhou, K., Lin, W., Sun, J., Zhang, J., Zhang, D., Feng, X., & Liu, Q. (2022). Prediction model of end-point phosphorus content for BOF based on monotone-constrained BP neural network. *Journal of Iron and Steel Research International*, *29*(5), 751–760. https://doi.org/10.1007/s42243-021-00655-6

[19] Zou, Y., Yang, L., Li, B., Yan, Z., Li, Z., Wang, S., & Guo, Y. (2022). Prediction model of end-point phosphorus content in EAF steelmaking based on BP neural network with periodical data optimization. *Metals*, *12*(9), 1519. https://doi.org/10.3390/met12091519

[20] Wang, H., Xu, A., Ai, L., & Tian, N. (2012). Prediction of endpoint phosphorus content of molten steel in BOF using weighted K-means and GMDH neural network. *Journal of Iron and Steel Research International*, *19*(1), 11–16. https://doi.org/10.1016/s1006-706x(12)60040-5

[21] Li, C. R., Zhao, H. W., & Yin, Q. (2011). Prediction model of end-point carbon content for BOF based on LM BP neural network. *Advanced Materials Research*, *189–193*, 4446–4450. https://doi.org/10.4028/www.scientific.net/amr.189-193.4446

[22] Liu, Z., Cheng, S., & Liu, P. (2022). Prediction model of BOF end-point P and O contents based on PCA–GA–BP neural network. *High Temperature Materials and Processes*, *41*(1), 505–513. https://doi.org/10.1515/htmp-2022-0050

[23] Cvetković, D., Šovljanski, O., Ranitović, A., Tomić, A., Markov, S., Savić, D., . . . , & Pezo, L. (2022). An artificial neural network as a tool for kombucha fermentation improvement. *Chemical Industry and Chemical Engineering Quarterly*, *28*(4), 277–286. https://doi.org/10.2298/ciceq211013002c

[24] Sorsa, A., Ruuska, J., Lilja, J., & Leiviskä, K. (2015). Data-driven multivariate analysis of basic oxygen furnace used in steel industry. *IFAC-PapersOnLine*, *48*(17), 177–182. https://doi.org/10.1016/j.ifacol.2015.10.099

[25] Niquini, G. R., Silva, S. R., Costa Junior, E. F., & Costa, A. O. S. (2019). Feedstock and inoculum characteristics and process parameters as predictors for methane yield in mesophilic solid-state anaerobic digestion. *Anais da Academia Brasileira de Ciências*, *91*(4), e20181181. https://doi.org/10.1590/0001-3765201920181181

How to Cite: Souza Chaves, D. H., Duarte, I. C. D., da Costa Junior, E. F., & da Costa, A. O. S. (2024). Comparison Between Empirical Strategies for Predicting Endpoint Phosphorus Content in BOF Steelmaking Process. *Archives of Advanced Engineering Science*. https://doi.org/10.47852/bonviewAAES42023358