

Forecasting Modeling for Energetic Efficiency in an Industrial Process

Simone M. Acosta^a, Angelo M. O. Sant'Anna^b, Osiris Canciglieri Junior^{c*}

^aDepartment of Electronics, Federal University of Technology Paraná (UTFPR), Av. Sete de Setembro, 3165, Curitiba, Brazil

^bMechanical Engineering Department, Federal University of Bahia, Rua Aristides Novis, 2, Salvador, Brazil.

^cIndustrial and Systems Engineering Graduate Program, Pontifical Catholic University of Parana, Rua Imaculada Conceição, 1155, Curitiba, Brazil.
osiris.canciglieri@pucpr.br

The quality of products and processes is a permanent challenge for industries, and such challenge is no different in steelmaking processes. One of the main problems affecting the quality of steel products is the existence of contaminants in alloy steel, and phosphorus (P) is a major contamination element interfering with the steelmaking process. The increased P levels can severely affect the physical integrity of steel bonds, thus threatening the quality of the final product. This paper proposes a robust approach to model the phosphorus concentration levels in the steelmaking process. The proposed approach consists in applying the artificial neural networks techniques for improving the energetic efficiency of the industrial process. We used the improved neural network models inspired in the human nervous system for processing the information. The different techniques used for modelling the phosphorus levels investigate the variables that have a significant influence on refining process. Based on the better predictive model, the increase of phosphorus levels in the final product is related to initial levels of carbon, oxygen, magnesium, manganese oxide and calcium oxide. The results illustrate the efficiency of the techniques used in the modelling, with emphasis on the adequacy of the predictive models constraints in the refining process. This study presents a relevant strategy to model characteristics' of raw material based on forecasting strategy to the efficiency of alloys and steel industry.

1. Introduction

Assuring the quality of processes and products remains a constant challenge for companies, and this is increasingly becoming a fundamental requirement for their endurance. This is no different in steelmaking companies. One of the main parameters affecting the excellence of products in this sector is the existence of contaminants in alloy steel. Phosphorus (P) is one of the main contaminants found to interfere in steelmaking processes. The ferromanganese alloys are the major sources of P contamination during the process of steelmaking, which demands a limited use of this kind of alloy along the process (Um et al., 2014). The increased P levels can significantly affect physical aspects of alloy steel, thus severely compromising its quality. P-rich steel compounds usually show: (i) increased hardness; (ii) decreased ductility; (iii) appearance of ghost lines in carbon-rich alloy steels; and (iv) increased frailty of steel bonds under high and low temperatures (Chaudhary et al., 2001).

The refining process adopted by the company subject to this study uses high-purity oxygen to reduce carbon levels in High-Carbon Ferromanganese (FeMnHC) originating Medium-Carbon Ferromanganese (FeMnMC) which has higher market value. During this process, there is an alteration in levels of other elements, including phosphorous. Therefore, the present work proposes a model for the FeMnMC refining process in an important Brazilian metallurgic industry with the purpose of predicting P levels at the end of the refining process. The Artificial Neural Network (ANN) method, defined as data processing systems based on the behavior of the human nervous system, was the chosen modeling technique (De Castro, 2007). The main advantages of ANNs are: (i) the capacity of approximating the behavior of non-linear physical phenomena, not requiring

profound statistical comprehension of the modeled data and; (ii) the capacity of learning any input/output variables of the continuous form (Haykin, 2008). An ANN is usually called a black-box model, considering that this kind of technique approximates the relative input/output of a certain process. For the process studied here, input variables were FeMnHC composition, dross composition, alkalinity indexes and amounts of solid loads added during the process. The output variable was P concentration at the end of the process.

Application of ANN to model different chemical and industrial processes is relevant and several authors have profited from the approximation of processes with complex behavior allowed by this technique in their work, such as Bakar et al. (2013) that performed a study to identify values of key parameters for prediction of heat exchanger of a bio-methanol production plant. Joo et al. (2013) propose to control the temperature of a polymerization reaction of styrene in suspension and to compare with classical controller PID systems. Osulale and Zhang (2014) evaluated different ANN models for forecasting hourly load of an electricity market. Yong et al. (2015) minimized the heat exchanger for waste heat under varying feed conditions using neural network model, allowing lower time and cost in manufacturing, while Gao and You (2015) developed an approach to optimize the pattern of filter networks, offering practitioners a relevant trade-off between quality and cost. This paper proposes a robust approach to model the phosphorus concentration levels in the steelmaking process. The proposed approach consists in applying the artificial neural networks techniques for improving the industrial process energetic efficiency.

2. Background

Artificial Neural Networks are systems that manipulate information through the interaction of basic processing units called artificial neurons (Haykin, 2008). Artificial neurons, in turn, are basic units that receive and process external inputs generating outputs. Du and Swamy (2013) asserted that an artificial neuron comprises the following components: (i) a group of synaptic weights that can balance the importance of input signals; (ii) a threshold term or Bias that enables to increase or decrease the impact of different inputs on neuronal function; (iii) an adder, which can add up values of inputs balanced according to synaptic weights and (iv) an activation function that controls the output signal. The artificial neural are usually organized in layers which can be categorized into output and hidden layers.

An ANN may have a variable number of hidden layers, but according to Kolmogorov's theorem and to the second theorem of universal approximation only one hidden layer is enough to approximate any function of continuous nature (Scarselli and Tsoi, 1998). This allows a reduction in the complexity of the model, but occurrence an increased number of neurons are required. An ANN model with two layers may be mathematically described by the Eq(1),

$$\hat{y}_t = f_o \left(\sum_{j=1}^m w_{kj}^o f_h \left(\sum_{i=1}^n w_{ji}^h x_t + b_j \right) + b_k \right); k = 1, 2, 3 \dots \quad (1)$$

where y_t is the estimated network output, w_{kj} and w_{ji} are the synaptic weights of the output and hidden layers, b_j represents the threshold terms, x_t is the values of input variables and $f(\cdot)$ symbolizes the exchange between neurons in output and hidden layers. Even though there are several forms of transfer functions, the most commonly studied are the sigmoid logistic and the linear and hyperbolic tangent functions.

The choice of the most suitable activation function is a crucial step and a linear function is typically used for the output layer in issues of function approximation, whereas logistic or tangent hyperbolic functions are employed in the hidden layers. Before the learning process starts, the data are standardized according to the activation function in the hidden layer. For logistic function, the data are standardized in the interval [0,1] whereas the interval [-1,1] is used for the hyperbolic tangent function. This standardization process is ruled by the Eq(2),

$$t_s = \frac{x_t - x_{min}}{x_{max} - x_{min}} (t_{max} - t_{min}) + t_{min} \quad (2)$$

where t_s is the standardized variable, t_{min} and t_{max} are the minimum and maximum values to t_s , respectively, x_t is the original variable, x_{min} and x_{max} are the minimum and maximum values of x_t .

These functions are chosen due to easy the derivation and simplification of the network learning process, at the synaptic weight regulation (Engelbracht, 2006). The learning process consists of two basic steps: (i) the prediction stage, where synaptic weights are adjusted based on a group of training data, and (ii) the test stage, that determines whether the prediction model is able to generalize.

The synaptic weights are set through an optimization algorithm called learning algorithm, which for multiple layer ANNs is usually the error Backpropagation algorithm. It consists on the minimization of the mean quadratic error (Eq(4)) of output variables generated by the ANN relative to real measured values.

There are several variations of the error Backpropagation algorithm, being the Levenberg-Marquadt algorithm amongst the most efficient. This algorithm, however, tends to cause overfitting, where the model specializes in

the training data. For that reason, a Bayesian regularization term can be employed in the ANN cost function. The ANN cost function was then modified as Eq(3),

$$\varepsilon = \alpha \left(\frac{1}{N} e_t^2 \right) + \beta \left(\sum_{j=0}^m w_{kj}^O + \sum_{i=0}^n w_{ji}^H \right) \quad (3)$$

where $e_t = (y_i - y_t)$ is the estimated error, y_i is the actual value measured at the system, y_t is the output estimated by the ANN, n is the number of samples, α and β are terms estimated by the learning algorithm with a cost function, w_{kj} and w_{ji} are the synaptic weights of output and hidden layers, respectively.

To evaluate the generalization capacity of an ANN error minimization strategies are used, such as the root mean square error (RMSE) and the mean absolute percentage error (MAPE), these can be determined according to Eq(4) and 5,

$$RMSE = \sqrt{\frac{1}{N} e_t^2} \quad (4)$$

$$MAPE = \frac{1}{N} \sum_{t=1}^n \left| \frac{e_t}{y_t} \right| \quad (5)$$

An important way for evaluation of an ANN model is to estimate the number of parameters or synaptic weights (W) and the amount of data used for model estimation (N), where corrected Akaike Information Criteria (AICc) is suitable (Sant'Anna, 2015), see Eq(6), and the number of synaptic weight in an ANN with two layers can be defined by the Eq(6).

$$AICc = \ln \left(\frac{\sum_{t=1}^N e_i}{N} \right) + \frac{2N(W + 1)}{N - W - 2} \quad (6)$$

where $W = [(k) + j] + [(j) + i]$, k represents the number of input variables, j is the number of neurons in the hidden layer and i the number of neurons in the output layer. (Vining and Kuwalski, 2011) assert that, an ANN model must have $N > W$ based on the Occam theorem that represents a system employing a minimum number of parameters and synaptic weights.

3. Case Study

Several kinds of steel are affected by high levels of phosphorus or manganese, which can compromise their quality. The refining process of ferromanganese consists of carbothermic reaction that involves the reduction and control of the percentage of carbon in the liquid metal by injection of oxygen in the bath.

The chamber for the refining process of ferromanganese on the oxygen converter has an internal layer covered of magnesium oxide (MgO), and two pipes (N₂-cooled) and vertical stream (water-cooled) blow oxygen. Inside the converter, the FeMnHC alloy has an average carbon level of 6.5 %, which increases the efficacy of the oxidant dephosphorylation. The Lime is also used during the dephosphorylation step, causing significant loss of manganese with no substantial effect on P levels. When lime is dissolved in the molten bath it increases the alkalinity of dross, which prevents the loss of manganese into the dross. Before the oxygen blow begins, approximately 60 kg of CaO per ton of FeMnHC dissolved in the bath are added, in a process divided into 3 stages.

On the first stage, dephosphorylation and decarburization start to occur and significant amounts of manganese are lost to the dross through volatilization, because of the excess of oxygen. During the second stage, called decarburization, bath temperature rises, reducing manganese oxidation and increasing the rate of decarburization as well as Mn volatilization. Rephosphorylation occurs as a consequence of such temperature variations because P₂O₅ is highly unstable at high temperatures. As the temperature continues to rise, rephosphorylation can only be halted if the activity of phosphor pentoxide is reduced, which can be achieved by dissolving P₂O₅ into the dross in the presence of Lime. At the third stage, there is the addition of a solid deoxidizing load containing 110 kg of HCFeMn or MCFeMn, 150 kg of FeSiMn and 40 kg of CaO per ton of liquid HCFeMn in the bath, which helps to cool the mixture and contributes to reducing manganese volatilization. At this point, a blow of inert CO₂ gas at 200 Nm³/h begins with the only purpose of stirring the bath.

3.1 Experimental design

The 21 initial input variables defined as relevant for modeling the dephosphorization process were grouped as: (i) composition of metal alloys (x_1, \dots, x_6); (ii) composition of slag (x_7, \dots, x_{15}); (iii) composition of solid loads

(x_{16} , x_{17} , x_{18}); and (iv) levels of alkalinity: binary; quaternary e optical basicity (x_{19} , x_{20} , x_{21}), see Table 1. The data selected covered a sample of 248 observations.

Table 1: Description of characteristics of sustainable products

Group	Specifications	Und
Alloys composition	Si*, P*, Fe*, Ti*, C*, O ₂	%
Slag composition	MgO, MnO, CaO, SiO ₂ , Al ₂ O ₃ , BaO, K ₂ O, TiO ₂ , FeO	wt-%
Loads composition	Initial, Liquid, Dross	Nm ³
Alkalinity levels	BB, BQ, BO	Nm ³

Considering the large number of analyzed variables, a correlation analysis was needed in order to select those which are actually linked to the P concentration levels at the final process. The Pearson correlation test is a statistical tool which indicates the correlation between two variables (to value $p < 0.01$). After the correlation test was applied it became clear that from the 21 initial variables only 4 (four) variables were significantly correlated with P level output variables at the final refining process, as the initial phosphorus level (P*), the initial carbon level (C*), the manganese oxide level (MnO) and the loads composition (liquid). The other variables present in the dephosphorization process haven't high influence to predict P concentration level at the final process. Table 2 presents the statistical analysis of outcome elements and P concentration level (output variable) and the Figure 1 illustrates the Phosphorus level data at the final refining process.

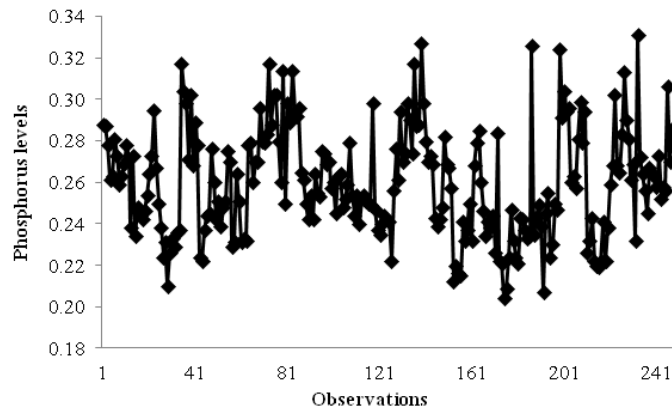


Figure 1: Phosphorus concentration level at the final process

Table 2: Summary of the variables selected

Group	Variable	Mean	Std. dev.	Max	Min	Correlation
Alloys	P*	0.216	0.025	0.280	0.155	0.687
Alloys	C*	6.759	0.140	7.070	6.077	-0.291
Slag	MnO	39.056	5.848	51.993	24.202	0.289
Loads	Liquid	9.889	1.543	12.700	5.510	-0.295
Output	P	0.260	0.026	0.331	0.204	1.000

The design of the proposed approach was performed using MATLAB®. The chosen output variable to the neural network was the P concentration level at the final Ferromanganese refining process. The relevant input variables selected were: initial phosphorus level (P*), initial carbon level (C*), manganese oxide level (MnO) and loads composition (Liquid).

After the pre-processing phase, it was randomly chosen 198 (80 %) observations of the 248 total observations performed for composing the Training group and 50 (20 %) were chosen for the Test group. It was used one hidden layer to projects in the ANN architecture, the Kolmogorov theorem, and the Levenberg-Marquadt algorithm. The Kolmogorov theorem found search-space for our study is 3.0 neurons which are in agreement with the established restriction: $N > W$ and $T = 198$ observations. The modeling process consisted of estimating ANN models with logistic and hyperbolic activation function in the hidden layer to assess which was the most efficient model, using 4 (four) input variables, one output variables, crossover rate 0.8, mutation rate 1.6,

learning rate 0.05, momentum rate 0.1 and population 198 obs. The appendices show the realized performance study for logistic and hyperbolic activation function (see Table 3).

The ANN model which better represented the process studied is with 11 neurons in the hidden layer with the logistic activation function, considering the error minimizing statistics and the complexity analysis of the model (see Tables 3). For 11 neurons in the hidden layer, the number of adjustable parameters is $W = 67$, keeping compliance to $N > W$. The Table 3 illustrates that as the number of neurons increases towards the top limit ($j_{max} = 14$) and the estimates of root mean square error (RMSE) and mean absolute percentage error (MAPE), highlights that the power of estimation from model decreases and its complexity increases.

The residuals errors generated by prediction model show a random distribution varying between - 0.04 and 0.05, without seasonal variation (see Figure 2). In general, the fitted ANN model showed a mean percentage error of 0.09 %, and it is possible to consider this ANN model as trustworthy for P concentration level at the final refining process.

Table 3: Performance of ANN with logistic activation function

Nodes	10	11	12	13	14
RMSE	0.0159	0.0159	0.0159	0.0159	0.0159
MAPE	0.000958	0.000949	0.000956	0.000955	0.000949
AICc	53.25	59.17	65.24	71.28	77.21

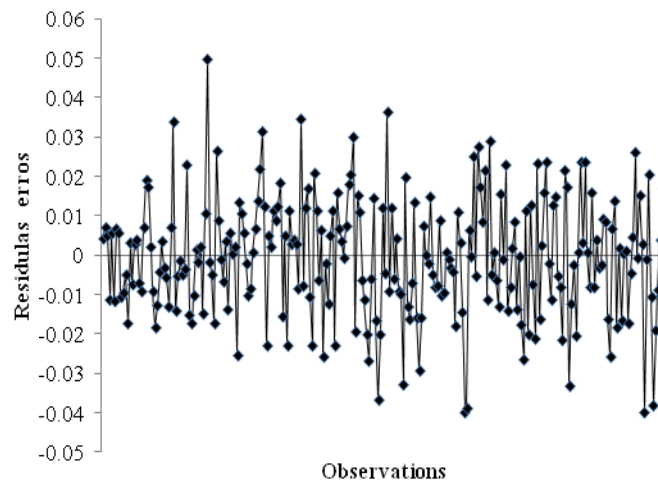


Figure 2: Residuals errors from the better ANN model

4. Conclusions

The process of extraction of metals, such as zinc, aluminum, iron and manganese for the production of alloys is very important to the industry, detaching the interest for reducing costs, environmental impacts and improvement performance. We propose to determine a prediction model for Ferromanganese refining process in a steelmaker industry that allows estimating the phosphorus concentration levels. The Ferromanganese refining process performed in a steelmaker industry to allow the prediction of P concentration levels. The development of the ANN model with the variables from Ferromanganese refining process was used as input and selected from correlation test. All the variables were tested with a correlation test to verify how each of them is related to the percentage of phosphorus in the final product. These results show that four variables were correlated and those were used to compose the input set of the ANN model.

The architecture of neural network was composed by the Kolmogorov theorem, the Levenberg-Marquardt algorithm, and diagnostic criteria to improve the parameters estimates, increase the quality of prediction and decrease the statistical errors. The samples were divided into groups: one was used to the training of the ANN model and a second was used to test the ANN model, those groups were composed of 198 and 50 samples, respectively. The great number of neurons in the hidden layers made the network structure very heavy in a computational way which demands a little more time to achieve the results, and the diagnostic analysis suggests 11 neurons, hyperbolic function to better performance of ANN model. The estimation model showed excellent results with an average percentage error of 0.09 %, supporting the generalization capacity and reliability of the ANN model. Based on this created estimation model it is possible to estimate the impact of certain P concentration levels in FeMnMC beforehand, with a considerable amount of reliability.

The present study has established that dephosphorization process of ferromanganese with alloys composition correlated variables: initial phosphorus (< 0.28 %), initial carbon (< 7.070 %) can be carried out to significant influence to estimate P concentration level. As soon, the manganese oxide level (24.20 – 51.993 wt.%) and loads liquid (5.51 – 12.7 Nm³) for slag and loads composition for dephosphorisation of high-carbon ferromanganese. To future research is possible to investigate non-linear or chaotic behavior between input and output variables, as an interesting point to be studied to make forecasting models.

Acknowledgments

The authors are grateful to the National Council of Scientific and Technological Development (CNPq) under grants 486707/2013-0 (Universal) and scholarship for the financial support of this work.

Reference

- Bakar S.H.A., Hamid M.K.A., Wan Alwi S.R., Manan Z.A., 2013, Flexible and Operable Heat Exchanger Networks, *Chemical Engineering Transactions*, 32, 1297–1302.
- Chaudhary P.N., Goel R.P., Roy G.G., 2001, Dephosphorisation of high carbon ferromanganese using BaCO₃ based fluxes, *Ironmaking and Steelmaking*, 28(5), 396–403.
- De Castro L.N., 2007, Fundamentals of natural computing: an overview, *Physics of Life Reviews*, 4, 1–36.
- Du K-L, Swamy M.N.S., 2013, *Neural networks and statistical learning*. Heidelberg: Springer.
- Engelbrecht A.P., 2006, *Computational Intelligence: An Introduction*. 2nd ed, West Sussex: John Wiley & Sons Ltd.
- Haykin S., 2008, *Neural networks and learning machine*. 3rd ed, New York: Prentice Hall.
- Joo S-W., Hong S-H., Lee G-H, You B-D., 2013, Evaporation behavior of low carbon ferromanganese alloy melt at reduced pressure, *Metal and Materials International*, 19(3), 585–590.
- Gao J., You F., 2015, Milfp Model and Algorithms for Network Design and Long-term Planning of Water Management System for Shale Gas Production, *Chemical Engineering Transactions*, 43, 1423–1428.
- Osolale F.N., Zhang J., 2014, Energy Efficient Control and Optimisation of Distillation Column Using Artificial Neural Network, *Chemical Engineering Transactions*, 39, 37–42.
- Sant'Anna A.M.O., 2015, Framework of decision in data modeling for quality improvement, *The TQM Journal*, 27(1), 135–149.
- Scarselli F., Tsoi A.C., 1998, Universal approximation using feed-forward neural networks: a survey of some existing methods and some new results, *Neural Networks*, 11(1), 15–37.
- Um H., Lee K., Kim K-Y., Shin G., Chung Y., 2014, Effect of carbon content of ferromanganese alloy on corrosion behaviour of MgO–C refractory, *Ironmaking and Steelmaking*, 41(1), 31–27.
- Vining G., Kuwalski S., 2011, *Statistical methods for Engineers*, 3rd ed, Boston: Cengage Learning.
- Yong J.Y., Nemet A., Varbanov P.S., Klemes J.J., Cucek L., Kravanja Z., Mantelli V., 2015, Heat Exchanger Network Modification for Waste Heat Utilisation Under Varying Feed Conditions, *Chemical Engineering Transactions*, 43, 1279–1284.