

## Forecast model for dephosphorization process of ferromanganese steels using artificial neural networks

Lee Vinagre Monteiro<sup>1</sup>, Angelo Marcio Oliveira Santana<sup>1,2</sup>

<sup>1</sup> Industrial & Systems Engineering Graduate Program, Pontifical Catholic University of Parana, Rua Imaculada Conceição, 1155, CEP: 90215-901, Curitiba, PR, Brasil.

<sup>2</sup> Polytechnic School, Federal University of Bahia, Rua prof. Aristides Novis, 02, CEP: 40210-630, Salvador, BA, Brasil. e-mail: leemonteiro@gmail.com, angelo.santanna@ufba.br

---

### ABSTRACT

One of the main problems affecting the quality of steel products is the existence of contaminants in alloy steel, being phosphorus (P) a major contamination element interfering with the steelmaking process. The increased P concentration levels can severely affect physical integrity of steel bonds, thus threatening the quality of the final product. The dephosphorization process of Ferromanganese consists by carbothermic reaction that involves the control of the manganese volatilization and reduction of manganese oxide in injection of oxygen. Therefore, we propose to forecast model for dephosphorization process of Ferromanganese steels in a steelmaker industry, that allows estimating the phosphorus concentration levels at the final refining process. We chose the artificial neural network models because it is computational models inspired in the human nervous system and an architecture of neural network with the Levenberg-Marquadt algorithm and Kolmogorov theorem for improving the estimation technique. The developed model presented excellent performance with a percentage error of 0.09%. 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.

**Keywords:** Ferromanganese alloys, dephosphorization process, Neural networks, Kolmogorov theorem.

---

### 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 creates a limitation on the use of this type of alloy throughout the process [1]. 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 [2]. 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 alteration in levels of other elements, including P. 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 [3]. The main advantages of ANNs are: (i) the capacity of approximating the behavior of non-linear physical phenomena, not requiring profound comprehension statistical and complex statistical treatment of the modeled data and; (ii) the capacity of learning any input/output variables of continuous form [4].

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 process studied here, input variables were FeMnHC

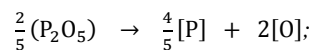
composition, dross composition, alkalinity indexes and amounts of solid loads added during the process. The output variable was P concentration at the end of the process. Application of ANN to model different chemical and industrial processes is relevant and several authors have profited from the approximation of processes with complex behavior allowed by this technique in their work, such as, it has been developed [5] a image system to quantify values of microstructures in steel and cast iron for volumetric fraction. In [6] propose to measure the acoustic emission for detection defects propagation in pipelines rigid. This method was performed to identify values of pressure, temperature and volume for evaporation of low carbon ferromanganese alloy, which occurs in alloy melt when increasing temperature and decreasing pressure [7] and [8] used ANN to model the Pleuromutilin fermentation process.

## 2. MATERIALS AND METHODS

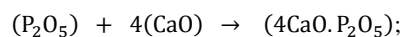
Several kinds of steel are affected by high levels of phosphorus or manganese, which can compromise their quality. The refining process of ferromanganese consists by carbothermic reaction that involves the reduction and control of the percentage of carbon in the liquid metal by injection of oxygen in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

The chamber for the refining process of ferromanganese on the oxygen converter has an internal layer covered of magnesium oxide (MgO), and two tuyère (N<sub>2</sub>-cooled) and a 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 dephosphorization. Lime is also used during the dephosphorization 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 60kg of CaO per ton of FeMnHC dissolved in the bath are added, in a process divided into 3 stages.

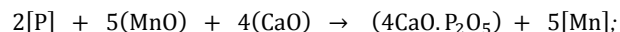
On the first stage, dephosphorization and decarburization start occurring and significant amounts of manganese are lost to the dross through volatilization because of the excess of oxygen. If there is still silicon content in the FeMnHC alloy to be reduced, the dephosphorization occurs at the final this stage following the silicon oxidation. During the second stage, called decarburization, bath temperature rises, reducing manganese oxidation and increasing the rate of decarburization as well as Mn volatilization. Rephosphorization occurs because of such temperature variations, because P<sub>2</sub>O<sub>5</sub> is highly unstable at high temperatures. As the temperature continues to rise, rephosphorization can only be halted if the activity of phosphor pentoxide is reduced, which can be achieved by dissolving P<sub>2</sub>O<sub>5</sub> into the dross in the presence of lime:

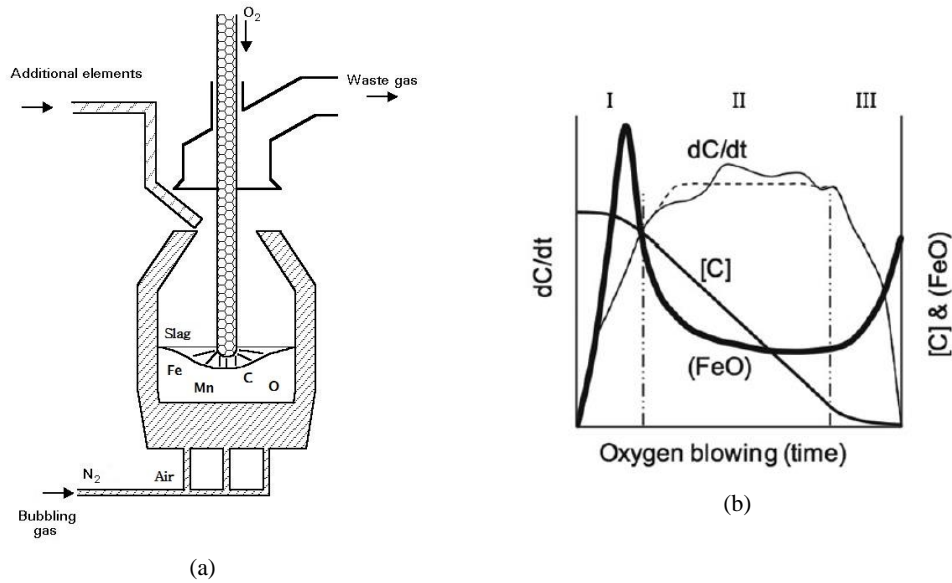


At the third stage, there is addition of a solid deoxidizing load containing 110kg of HCFeMn or MCFeMn, 150 kg of FeSiMn and 40 kg of CaO per ton of liquid HCFeMn in the bath, which helps cooling the mixture and contributes to reduce manganese volatilization. At this point, a blow of inert CO<sub>2</sub> gas at 200 Nm<sub>3</sub>/h begins with the only purpose of stirring the bath.



Due to the addition of lime in the converter, in this phase also occurs phosphorus oxidation of FeMnMC and reduction of manganese oxide of the slag. The Figure 1 illustrate the converter of oxygen with the vertical stream of flow rate of 350 Nm<sub>3</sub>/h in the upper part and two pipes with 800 Nm<sub>3</sub>/h in the lower part.

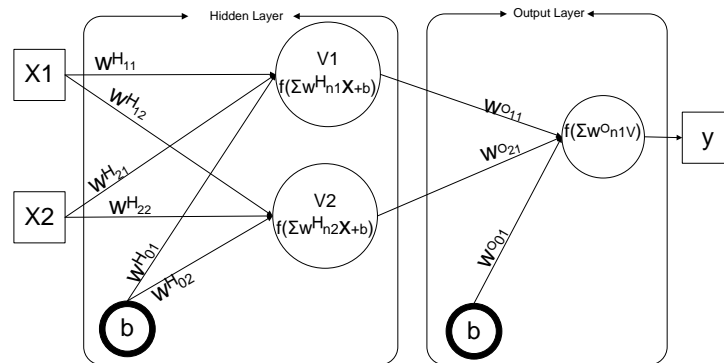




**Figure 1:** Oxygen converter of the studied process (a) and oxygen effects during dephosphorization process of steel (b)

**2.1 Artificial Neural Network (ANN)**

Artificial Neural Networks are systems that manipulate information through the interaction of basic processing units called artificial neurons [4]. Artificial neurons, in turn, are basic units that receive and process external inputs generating outputs. Figure 2 illustrates a conceptual model of ANN. In [9], 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.



**Figure 2:** Conceptual model of ANN.

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 [10]. This allows a reduction in the complexity of the model; but as a consequence, an increased number of neurons are required. An ANN with two layers may be mathematically described by the equations,

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

where  $\hat{y}_i$  is the estimated network output,  $w_{kj}^o$  and  $w_{ji}^h$  are the synaptic weights of the output and hidden layers, b represents the threshold terms,  $x_t$  are 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; 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, data are standardized according to the activation function in the hidden layer. For logistic function, 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 equation,

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

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

These functions are chosen due to easy to derivatives and simplify for network learning process, at the synaptic weight regulation [11]. 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 then on model is able to generalize. 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. 3) of output variables generated by the ANN relative to real measured values.

$$\varepsilon = \frac{(y_t - \hat{y}_t)^2}{N} \quad (3)$$

where  $\varepsilon_t$  is the estimated error,  $y_t$  is the actual value measured at the system,  $\hat{y}_t$  is the output estimated by the ANN and  $N$  is the number of samples.

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, it was proposed [12] a Bayesian regularization term be employed in the ANN cost function. The ANN cost function was then modified as,

$$\varepsilon = \alpha \left( \frac{1}{N} e_t^2 \right) + \beta \left( \sum_{j=0}^m w_{kj}^o + \sum_{i=0}^n w_{ji}^h \right) \quad (4)$$

where  $\alpha$  and  $\beta$  are terms estimated by the learning algorithm with a cost function,  $w_{kj}^o$  and  $w_{ji}^h$  are the synaptic weights of output and hidden layers, respectively. To evaluate the generalization capacity of an ANN error minimization and model estimation strategies are used, such as: the root mean square error (RMSE) and the mean absolute percentage error (MAPE) and corrected Akaike Information Criteria (AICc),

$$RMSE = \sqrt{\frac{1}{N} \varepsilon_t^2} \quad \text{and} \quad MAPE = \frac{1}{N} \sum_{t=1}^n \left| \frac{\varepsilon_t}{y_t} \right| \quad (5)$$

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

### 3. RESULTS

The 21 initial input variables defined as relevant for modeling the dephosphorisation 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.

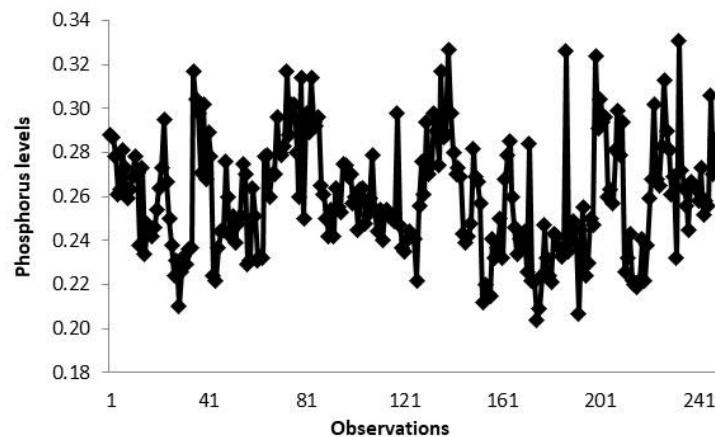
**Table 1:** Variables in the studied process.

GROUP	VARIABLE	UND
Raw Materials	Si*, P*, Fe*, Ti*, C*, O <sub>2</sub>	%
Dross composition	MgO, MnO, CaO, SiO <sub>2</sub> , Al <sub>2</sub> O <sub>3</sub> , BaO, K <sub>2</sub> O, TiO <sub>2</sub> , FeO	wt-%
Volume composition	Initial, Liquid, Dross	Nm <sup>3</sup>
Alkalinity levels	BB, BQ, BO	Nm <sup>3</sup>

Considering the large number of variables analyzed, 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  $p < 0.01$ ) [13]. 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. The Table 2 presents the statistical analysis of outcome elements and P concentration level (output variable) and the Figure 3 illustrates the Phosphorus level data from refining process.

**Table 2:** Statistical summary of the variables selected.

GROUP	VARIABLES	AVERAGE	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



**Figure 3:** Phosphorus concentration at the final process

The design of the proposed ANN and statistical analysis were performed using Matlab®. The chosen output variable to neural network was the P concentration level at the end Ferromanganese refining process. The input variables were: initial phosphorus content (P\*), initial carbon content (C\*), manganese oxide content (MnO) and Volume Composition (Liquid).

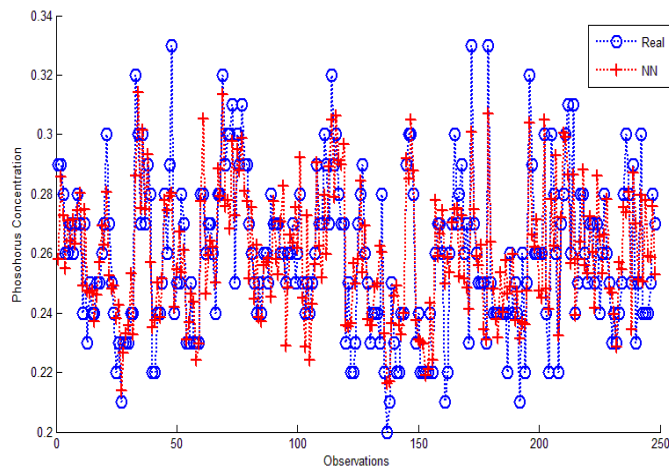
After the pre-processing phase, it was randomly chosen 198 (80%) observations of the 248 total observations performed for composing the Training group and 50 (20%) were chosen for the Test group. It was used one hidden layer to projects in the ANN architecture, the Kolmogorov theorem and the Levenberg-Marquadt algorithm. The Kolmogorov theorem found search-space for our study with neurons which is in agreement with the established restriction:  $N > w$  and  $T=198$  observations [10]. The modelling process consisted to estimate ANN models with logistic and hyperbolic activation function in the hidden layer to assess which was more 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 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 Table 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 ( $j_{max} = 14$ ), the power of estimation from model decreases and its complexity increases.

**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

The residuals or prediction errors generated by prediction model show a random distribution varying between -0.04 and 0.05, with no seasonal variation or tendencies ( $p < 0.01$ ) [14]. In general, the prediction regression 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. The Figure 2 illustrates a comparative between real data and predicted values of P concentration level at the final FeMnMC refinement process.

**Figure 4:** Comparison of Phosphorus concentration values real and predicted after modeling.

#### 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 were 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 with 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: 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 in groups: one used to the training of the ANN model and a second used to test the ANN model, those groups were composed by 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 dephosphorisation 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 – 51.993wt-%) and loads liquid (5.51 – 12.7Nm<sup>3</sup>) for slag and loads composition for dephosphorisation of high carbon ferromanganese. To future research is possible to investigate non-linear or chaotic behaviour between input and output variables, as interesting point to be studied to make forecasting models.

## 5. ACKNOWLEDGMENTS

The authors are thankful for the financial support provided by National Council of Scientific and Technologic Development of Brazil (CNPq) under grants 486707/2013-0 and we gratefully acknowledge the anonymous referees for their valuable comments and suggestions.

## 6. BIBLIOGRAPHY

- [1] UM, H., LEE, K., KIM, K-Y., *et al.*, “Effect of carbon content of ferromanganese alloy on corrosion behavior of MgO–C refractory”, *Ironmaking & Steelmaking*, v. 41, n. 1, pp. 31–27, 2014.
- [2] CHAUDHARY, P.N., GOEL, R.P., ROY, G.G., “Dephosphorization of high carbon ferromanganese using BaCO<sub>3</sub> based fluxes”, *Ironmaking & Steelmaking*, v. 28, n. 5, pp. 396–403, 2001.
- [3] CASTRO, L.N., “Fundamentals of natural computing: an overview”, *Physics of Life Reviews*, v. 4, n. 1, pp. 1–36, 2007.
- [4] HAYKIN, S., *Neural networks and learning machine*, 3<sup>rd</sup> ed., New York, Prentice Hall, 2008.
- [5] ALBUQUERQUE, V.H.C., CORTEZ, P.C., ALEXANDRIA, A.R., *et al.*, “Image segmentation system for quantification of microstructures in metals using artificial neural networks”, *Revista Matéria*, v. 12, n. 2, pp. 394–407, 2007.
- [6] PINTO, C.F.C., SILVA, R.R., CALÔBA, L.P., *et al.*, “Use of artificial neural networks in detection of propagation of defects in pipelines rigid”, *Revista Matéria*, v. 17, n. 3, pp. 1084–1097, 2012.
- [7] JOO, S-W., HONG, S-H., LEE, G-H., YOU, B-D., “Evaporation behavior of low carbon ferromanganese alloy melt at reduced pressure”, *Metal and Materials International*, v. 19, n. 3, pp. 585–590, 2013.
- [8] KHAOUANE, L., BENKORTBI, O., HANINI, S., *et al.*, “Modeling of an industrial process of pleuromutilin fermentation using feedforward neural networks”, *Brazilian Journal of Chemical Engineering*, v. 30, n. 1, pp. 105–116, 2013.
- [9] DU, K-L., SWAMY, M.N.S., *Neural networks and statistical learning*, Heidelberg, Springer, 2013.
- [10] SCARSELLI, F., TSOI, A.C., “Universal approximation using feedforward neural networks: a survey of some existing methods and some new results”, *Neural Networks*, v. 11, n. 1, pp. 15–37, 1998.
- [11] ENGELBRECHT, A.P., *Computational Intelligence: an introduction*, 2<sup>nd</sup> ed., West Sussex, JohnWiley & Sons, 2006.
- [12] MACKAY, J.C.D., “Bayesian interpolation”, *Neural Computation*, v. 4, n. 2, pp. 415–447, 1992.
- [13] SANT’ANNA, A.M.O., CATEN, C.S., “Modelling of the nonconformig fraction in industrial processes”, *Pesquisa Operacional*, v. 30, n. 1, pp. 53–72, 2010.
- [14] VINING, G., KUWALSKI, S., *Statistical methods for engineers*, 3<sup>rd</sup> ed., Boston, Cengage Learning, 2011.

## ORCID

Lee Vinagre Monteiro

<http://orcid.org/0000-0002-7883-9450>

Angelo Marcio Oliveira Santanna

<http://orcid.org/0000-0001-8332-8877>