

APPLICATION OF NEURAL NETWORKS FOR MODELING STEELMAKING PROCESS

Lee Vinagre Monteiro

Industrial and System Engineering Graduate Program/PUCPR
Imaculada Conceição St.1155. Curitiba/PR
leemonteiro14@gmail.com

Ângelo Marcio Oliveira Sant'Anna

Industrial and System Engineering Graduate Program/PUCPR
Imaculada Conceição St.1155. Curitiba/PR
angelo.santanna@puc.br

ABSTRACT

One of the major causes of unconformity in the steelmaking process is the presence of impurities in the steel alloys. One of the main impurities found in the steel alloys is the phosphorus. The phosphorus causes several deleterious effects in the steel. In that way this research has as main objective to predict the percentage of phosphorus in the final composition of the material in the production process of one chemical industry. Our approach used one model of Artificial Neural Network in the data of the production process disposed. After apply the network it was compared with regression models developed with the same objective using statistical measures of error minimization. The results present that the network obtained the best results, an error minimization of 3, 68%.

Keywords: Artificial Neural Networks, Steelmaking process, Percentage of Phosphorus.

1. Introduction

Nowadays ensure the quality of products is a constant challenge. The main reason that justifies this is the quality, a basic survival requirement of a company which it is required to control and assure the quality products. However there are many factors that hamper these practices.

That challenge is not different in steelmaking companies, one of the factors that difficult the conformity of the products is the presence of impurities in several kinds of steel alloys. The phosphorus is the main impurity found in steelmaking process of medium Carbon Iron Manganese (FeMnMC) indispensable for the conversion of “pig iron” in steel.

According to Cruz (1994), Liu(1996) and Chaudary (2001) high levels of phosphorus in the composition of steel causes several effects in its physical characteristics like: (i) improvement of the hardness of steel; (ii) degree its ductility;(iii) it provides appearance of ghost lines in steel with high levels of carbon;(iv) improvement the fragility in high temperature and low ones.

Trying to solve this problem, Pedrini and Caten (2010) propose to use regression models for predicting the percentage of phosphorus in steelmaking process of Medium Carbon Iron Manganese (FeMnMC). With the application of the regression models they were able to found seven models that could predict the percentage of phosphorus with a considerable level of accuracy.

With that in mind, this research aims to apply Artificial Neural Networks (ANN) to predict the percentage of phosphorus in the composition of Medium Carbon Iron Manganese (FeMnMC). Applying the ANN, it was used as inputs, variables related with the steelmaking process of FeMnMC, particularly the percentage of elements used in its composition. The predictions of the percentage of phosphorus were considered as outputs variable. To verify the accuracy of the ANN, it was compared with the regression models.

In that way was formulated the main question of this research: Is it possible to develop an Artificial Neural Network that could predict the percentage of phosphorus with a considerable level of accuracy?

The application of ANNs to predict events is not new one. A considerable group of authors used ANNs in a large way to predict the most variables kinds of problems. Santos (2008) and Kirsten (2009) applied neural networks to predict temporal series of the financial market. Ghiassi and Nangoy (2009) used neural networks to predict nonlinear processes. Lee et al., (2009) employed neural networks to predict the Melt Index in an industrial process of ethylene and vinyl acetate. Myllykoski et al., (1996) used ANN to predict the mechanical properties of thin steel strip, a similar approach used by Perzyk and Kochansk (2001) where was used ANN to predict the ductility of iron. Col et al., (2007) applied a neural network to define the best machinery setup in the production of microalloyed steel. Eksioglu et al., (1996) applied neural networks in comparison with regression models in an ergonomic research in one industrial facility. All those authors had achieved good results applying neural networks.

The prediction is a good way to try to reduce the unnecessary spent with raw material and to avoid the unconformity in the production. Knowing the approximated quantity of raw material to be used and which they could cause collateral effects in the production is easier try to prevent unconformity.

2. Theoretical Review

In this section will be introduced the theoretical framework of this research which is divided in two parts: (i) the basic concepts of Artificial Neural Networks; (ii) Statistical tools that helped in the data analysis and interpretation of the results.

2.1. Artificial Neural Networks (ANN)

Haykin (2001) relates that ANNs are systems that manipulate information by the interaction of basic units of processing as known as artificial neurons. The system receives external signs, those are considered as inputs which are processed and then generate outputs. The

neurons send signals to each other and consequently the information will be transmitted to the whole network.

Artificial Neurons are basic units of processing in an ANN, they also are the fundamentals units to the operation of an ANN without artificial neurons a neural network cannot exist (HAYKIN, 2001). According to Haykin (2001), Braga (2007) and Kovacs (2004) one artificial neuron is composed of three basic elements: (i) One synapsis set or connection links which are characterized by one weight, these weights can assumed positive or negative values;(ii) One adder to add the input signals weighted by the respective neuron synapsis; (iii) One transfer function that restrict the amplitude neuron output. Depending on the type of ANN, the bias element could be present or not. The bias is a synapsis connection with unitary weigh. The Figure 1 shows a representation of an artificial neuron:

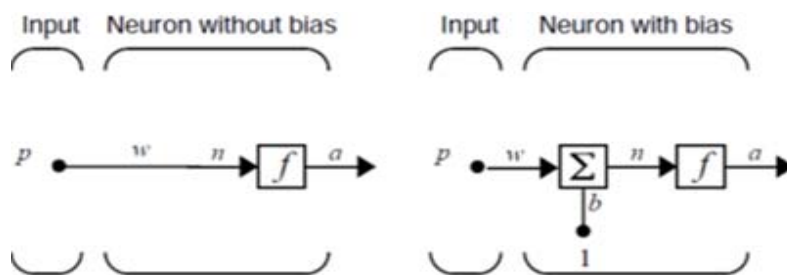


Figure 1 – Artificial Neuron Model
 Source: Demuth and Beale (2000)

The number of neurons of an ANN varies in accordance with the kind of problem that is wanted to solve. The number of neurons and layers (the way the neurons are organized) is defined in a heuristic way. The transfer function depends on the kind of ANN is used, normally are used the logistic functions, hyperbolic, tangent or linear ones.

2.1.1. Learning and Training

To apply ANNs it is necessary that the network goes through two stages: the stage of learning or training and the stage test. After those stages the ANN will be able to solve the problems which have been developed. ANNs are characterized by the ability of learning by examples; those are used to learn the ANN behavior of the phenomenon in study (BRAGA, 2007).

The training process or learning is the first step for application of an ANN to a specific problem. This learning process can be supervised or unsupervised process. The supervised training process consist in a presence of a supervisor that is responsible to excite the network inputs and observes the outputs found by the ANN with the desirables outputs in the test phase (BRAGA ,2007). The most common type of learning is the error correction learning models.

2.1.2. Multilayer Perceptron (MLP)

An ANN architecture is the form of neurons are organized. It contributes significantly in the network performance in several applications. According to Braga (2007), the ANN architecture can be categorized by the number of layers, by the connectivity and by the type of feed.

In this study was used one specific kind of ANN: the Multilayer Perceptron (MLP). The MLP can be defining as a multilayer network, totally connected (means that every neuron in one layer is connected with the next layer and so on). MLP networks are trained with the Back propagation Algorithm, a generalization of Delta rule. Delta rule is used to train ANN of only one layer. That kind of learning rule is one of error correction kind, so the learning process of MLP is the supervised type. The Figure 2 shows a model of MLP networks:

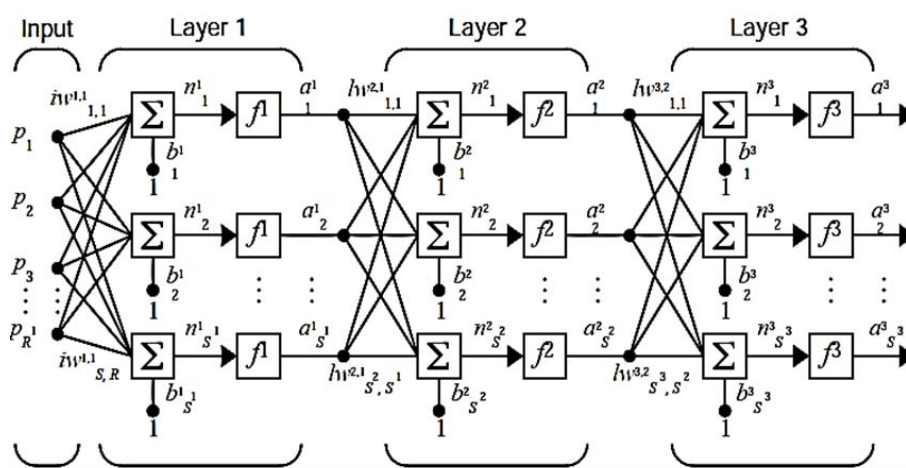


Figure 2 – MLP Model
 Source: Demuth and Beale (2000)

The Back Propagation algorithm was originally developed by Rumelhart, Hinton and Williams solving one problem that limited the training of complex networks. The major difficulty to the training process of ANN with multiple layers was the error that was obtained by the difference between the desirable outputs and generated outputs because there did not exist a desirable output for the intermediary layers (BRAGA, 2007; KOVACS, 2004). The Retro propagation algorithm works basically in two stages: the stage of propagation or forward and the stage of back propagation or backward.

In the propagation one input is apply to the network and this valor is propagated through the network until an output be produced. During the propagation the synapsis weights are fixed. This phase is used to define the network output by a defined input pattern. During the forward phase the training vector is train and is during this stage are made the calculus of the error correction rule. In the propagation one input is apply to the network and this valor is propagated through the network until an output be produced. During the propagation the synapsis weights are fixed. This phase is used to define the network output by a defined input pattern. During the forward phase the training vector is train and is during this stage are made the calculus of the error correction rule.

In the back propagation phase the synapsis weights are all adjusted according with the forward phase results. The answer produced by the network during this is used to recalculate the error correction rule and the synapsis weights are adjusted again. This process will be repeated over and over again until the specific value to be found. This cycle is called Epoch, and it's defined in the specifications of the ANN.

2.2. Statistical Tools

2.2.1. Correlation Analysis

Correlation Analysis is a statistical term used to design the force between one, two or more variables. Favero (2009) relates that the correlation is one of the most important techniques to trying to explain the nature of the relation between variables in study. This technique uses samples data analyses to obtain the information that could prove the existence or note the relation between the variables in study. The coefficient generated through the correlation analyses is the base to estimate various types of statistical relations like regression models

The coefficient of correlation value varies between -1 and 1. When the value of the coefficient is equal to -1 it is said that the correlation is perfectly negative when it is equal to 1 it is said that the correlation is perfectly positive. When the coefficient of correlation is zero it is said that the variables are not correlated. The correlation is considered significant if it value is bigger than 0, 05 or - 0, 05.

2.2.2. Statistics Measures of Error Minimization

Lustosa (2008) explains that the basic indicator to measure the prediction error for a determinate period of time t (e_t) is the difference between the real value (Y_t) and the predicted value (F_t) in the corresponding time. The formulation follows:

$$e_t = Y_t - F_t \quad (1)$$

The deviations from the "n" consecutive periods are calculated by the following measures: Mean Error (ME), Mean Absolute Error (MAE) and Mean Square Error (MSE). Another important measure for the performance analysis of the predictive models is the Absolute Percentage Error Medium (MAPE). The lower the values of these measures are, smaller will be the error obtained by the models used. These measures can be calculated by the following expressions:

$$EM = \frac{1}{n} \sum_{t=1}^n e_t \quad MAE = \frac{1}{n} \sum_{t=1}^n |e_t| \quad MSE = \frac{1}{n} \sum_{t=1}^n (e_t)^2 \quad MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right| \quad (2)$$

3. Case Study

The refine process of the chemical industry in study uses oxygen with high levels of purity to realize the decarburization process of FeMnAC producing FeMnMC which is much valorized by the market. During this process occur modifications in several elements percentage including between them the percentages of phosphorus. Pedrini and Caten (2010) discuss that the process of refine of "iron manganese" alloys consists in the reaction of decarburization between the liquid metal with the oxygen inject in the metallic bath. For realizing the process of dephosphorization is dissolved CaO during the decarburization in intent to reduce the percentage of phosphorus in the final product.

4. Materials and Methods

This section was divided in three parts: (i) the first part of this section gives an explanation of the database and the variables in study; (ii) the second part of this section explains how was made the data preprocessing to develop an ANN; (iii) the third part of this section explains how was the development of the ANN. To manipulate the database was used the software MS Excel® and to develop the ANN and apply the correlation tests was used the software MATLAB.

4.1. Database

In order to develop this research was used a database originally from the information system of the company in study. This database contains all the variables related with the steelmaking process of FeMnMC. Those variables are divided in five large groups: the first group is the initial composition; this group contains all the raw material used in the process. The second group contains all the scoria composition. The third group contains the volume composition of the material. The fourth group contains the basicity levels and the fifth group contains the final product composition. The variables used were the following:

Table 1 – Process variables investigated

Group	Variable
Raw Material	Si*, P*, Fe*, Ti*, C*, O2
Scoria composition	MgO, MnO, CaO, SiO2, Al2O3, BaO, K2O, TiO2, FeO
Volume composition	Initial, Liquid, Scoria
Basicity levels	BB, BQ, BO

The database is composed by 274 sample sets. Analyzing the database was discovered some missing values, those missing value were fulfilled with the average of that specific variable. It was made in order to prevent errors in the statistical tests. After the tests those sets were removed from the architecture of the ANN.

4.2. Preprocessing Phase

The preprocessing phase was the phase that provides all the data analyses the responsibility to the ANN development. In the first step was applied a correlation test to determinate which variable select to build the input set. The second step was a variable patterning. It means the variable selected was standardized to fit in an interval between 1 and 0. The importance of this will be explained later.

4.2.1. Correlation Test

The Pearson's Correlation test was used to select the variable to compose the input set. This test was applied in each one of the variables to verify the correlation of each one of them with the percentage variable of phosphorus. To compose the input set were selected the variables with the strongest levels of correlation

The variable selected were the following: Initial volume of the alloy, liquid volume of the alloy, percentage of BaO in the scoria, percentage of CaO in the scoria, percentage of SiO₂ in the scoria, percentage of MnO in the scoria, percentage of phosphorus in the raw material (P*), percentage of carbon in the raw material (C*) and the percentage of oxygen in the raw material (O₂). Those variables had following results in the correlation test:

Table 2 – Correlation coefficients to input set

Variable	Correlation
Initial	-0,2994
Liquid	-0,3055
BaO	-0,1576
CaO	-0,2737
SiO ₂	-0,1478
MnO	0,3002
P*	0,6821
C*	-0,3044
O ₂	-0,2240

4.2.2. Variable Patterning

To make the data processing easier to the ANN was applied a pattern equation in the variables values, this was necessary to make the network lighter. Depending on the values of the variables the network weights very large or very low to make the input values to fit in the transfer function. This action helps to make the processing become easier to the ANN learning. The equation used was the following:

$$X_p = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (3)$$

where: X_p is the standardized value; X_i is the observed value; X_{min} is the minimum value; X_{max} is the maximum value.

After apply this equation, the data were divided in two groups one composed by 240 samples and another smaller composed by 30 samples. These two groups were the training group and the test group respectively. These groups were selected in a random way. The numbers of samples with missing values were removed from the total of samples used in that data analyses.

4.3. ANN development

To develop the ANN was used the software MATLAB that has specific modules to develop ANN and others techniques of Computational Intelligence. With the input sets elaborated and all data adjusted was only a question to configure the parameters of the ANN in MATLAB neural tool box. After several tests, remind that the neural parameters are defined in heuristic way, was found the following configuration that presented the best results:

- Transfer Function: the transfer function selected to the ANN was the logistic function and was the unique type of function used in this ANN;
- Neurons: The total of neurons used was 100 distributed in three layers. The neurons used were sixty in the first layer, third nine in the second and one in the final layer;
- Error Minimization: was used an error minimization objective of 0,0001;
- Learning rate: was used a learning rate of 0,1;
- Number of Epochs: was used a number of 20.000 epochs.

With the parameters configured the network was able to be trained with the training group. After the training process the network was used to simulate the outputs of the test group and compared with the real value of the real observations. This network that was called “Net S4” shows the best results however it needs a major number of neurons and a large number of epochs to achieve those results.

5. Results

This research has as main objective to develop an Artificial Neural Network to predict the percentage of phosphorus in a specific steel alloy (FeMnMC). To achieve this goal was used a database with several variables related with the steelmaking process of FeMnMC.

Those variables were evaluated with Pearson’s correlation test to create the input set of the ANN. This input was formed by the variables: Initial volume of the alloy, Liquid volume of the alloy, percentage of BaO in the scoria, percentage of CaO in the scoria, percentage of SiO₂ in the scoria, percentage of MnO in the scoria, percentage of phosphorus in the raw material (P*), percentage of carbon in the raw material (C*) and the percentage of oxygen in the raw material (O₂).

The variables were used to create two groups: one training group composed by 240 samples and a test group composed by 30 samples. Those groups were elaborated in a random way. Before that all the values in the samples were standardized to fit in an interval between 1 and 0. The test group was used to train an MLP network trained with the Back propagation algorithm. After several tests one specific configuration of network was found and apply to create the network “Net S4”.

After the training phase of the ANN Net S4 the network was used to simulate the outputs of the test group. The same test group was simulated with the regression models developed by Pedrini and Caten (2010). All the simulations were used as a base to compare the models and to do it were used the statistical measures presented earlier and were found the following results:

Table 3 – Statistical measures of error minimization results

Models	ME	MAE	MSE	MAPE
Eq1	-0,003340289	0,015274352	0,000332116	5,75%
Eq2	0,001937275	0,014651036	0,000326946	5,38%
Eq4	0,002094556	0,014626624	0,000327177	5,36%
Eq5	0,001487449	0,014297187	0,000315444	5,24%
Eq6	0,000955160	0,015082436	0,000359652	5,52%
Eq8	-0,001665120	0,012851347	0,000249415	4,83%
Net S4	0,001180000	0,010293333	0,000233469	3,68%

Analyzing the results can be seen that the ANN had a better performance in the prediction of the percentage of phosphorus than the regression models. Comparing with each model the ANN Net S4 had a superior performance losing only to the equation 6 in the mean error (ME) measure. The comparison between the MAPE statistics could be better seen in the graphic bellow:

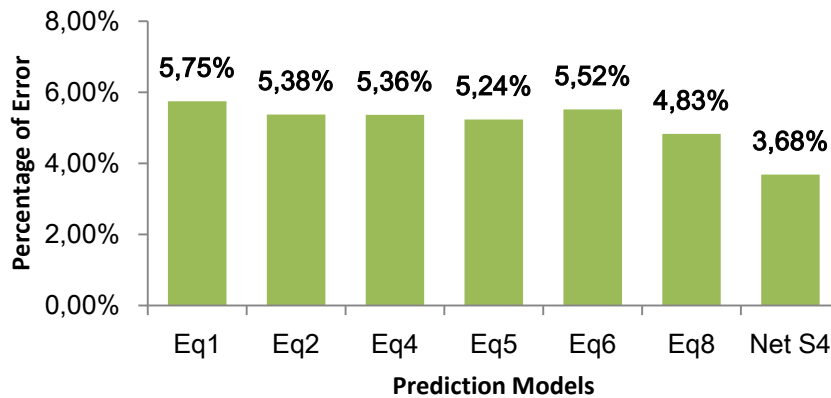


Figure 3 – MAPE comparison

Analyzing the results is possible to see that the equation 1 and the equation 8 present a large number of predictions with highest values than the real ones. According to Pedrini and Caten (2010) the equation 8 was the best model found to predict the percentage of phosphorus what was proved by the statistical analyzes. The equation 8 had the best results compared with the other equations losing only to equation 6 in the mean error (ME) measure.

Comparing the structure of the neural network with the regression models the Net S4 needed a larger number of variables. The models needed only three or four variables to predict the percentage of phosphorus while the Net S4 needed nine variables to predict the same value.

One important point related with the variables used is that almost all the variables used in the models were used to compose the ANN input set. Another important point related with the variables was that the percentage of Iron (Fe*) did not present correlation levels strong enough to make part of the input set. It is a very important point to be studied since all the models of regression used the percentage of Iron to make their prediction.

The use of more variables could be the main reason of the better performance of the ANN in comparison with the regression models and it helps to see that the percentage of phosphorus is influenced by a major number of factors than was considered before.

6. Summary and Conclusion

This research wanted to develop an artificial neural network that could be able to predict the percentage of phosphorus in the final constitution of “iron manganese” alloys, which could improve the steelmaking process decreasing the effects caused by the influence of high percentages of phosphorus in the main composition of the steel alloy.

The artificial neural network proposed was used in a database provided by the company in study that contains several variables related with the steelmaking process of the alloy. The database contains 274 samples that were used to all data analyzes and to develop the ANN. 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. In the end of these tests nine variables were selected and those were used to compose the input set of the ANN.

With the input set elaborated was needed to pattern the input set in order to improve the learning capability of the ANN. The samples were divided in to groups: one used to the training of the ANN and a second one used to test the ANN, those groups were composed by 240 and 30 samples respectively. With this was a question to apply the ANN with the data selected. After

several tests were found a configuration able to predict the phosphorus percentage, this network was called “Net S4”. In order to verify the reliability of the neural model found, the network was compared with the regression model proposed by Pedrini and Caten (2010). The models simulated the same samples used the test group. To made those comparisons were used some statistical measures of error minimization. Comparing the models was found the following conclusion: The ANN model achieve the best results in the prediction of the phosphorus percentage compared with the regression models with Absolute Percentage Error Medium of 3,8% and had superior measures in the other 2(two) measures.

With that comparison was able to conjecture the following conclusion:

- One neural network is able to predict the percentage of phosphorus with a consider level of accuracy. With that conclusion the main objective of this research was achieved;
- The neural model is a valid alternative to regression models;
- Comparing the variables in the neural model with the variables in the regression models was seen that the percentage of Iron (Fe*) is not an important variable to predict the percentage of phosphorus;
- The number of variables related with the percentage of phosphorus is bigger than was proposed in the regression models.

One negative point in the application of the ANN was its architecture. The parameters found were not the best way to achieve the main objective. The great number of neurons in the hidden layers made the network structure very heavy in a computational way, that and the elevated number of epochs to train the network made the process of application very slow what demand a little more time to achieve the results.

An alternative proposed to future researches is an application of other training algorithms like the Quasi-Newton or the Levenberg-Marquardt training algorithms. Another possibility is applying genetic algorithms or other techniques of Computational Intelligence to optimize the Neural Network configuration.

REFERENCES

- Braga, A.P.; Carvalho, A.F.; Ludermir, T.B.**. Redes Neurais Artificiais: teorias e aplicações. 2.ed. Rio de Janeiro: LTC, 2007.
- Corrar, L.J.; Paulo, E.; Filho, J. M.** Análise Multivariada: para os cursos de administração, ciências contábeis e economia. 1 ed. São Paulo: Atlas, 2007.
- Cruz, E.B.; Neto, J.B.F.; Yamada, M.; Monteiro, S.; Souza, J.O.** Desfosforação de Ferro Manganês. São Paulo: 26º Seminário sobre Fusão, Refino e Solidificação dos aço, p. 513-529, 1994.
- Col M., Ertunc H.M., Yilmaz M.** An artificial neural network model for toughness properties in microalloyed steel in consideration of industrial production conditions. *Materials and Design*, 28 (2), p. 488-495, 2007
- Chaudhary, P.N.; Goel, R.P.; Roy, G.G.** Dephosphorization of high carbon ferromanganese using BaCO₃ based fluxes. *Ironmaking and Steelmaking*, v.28, n. 5, p. 396-403, 2001.
- Demuth, H.; Beale, M.** Neural Network Toolbox for use with MATLAB (User’s Guide), 6 Ed. The Math Works, inc. 2000.
- Eksioglu M., Fernandez J.E., Twomey J.M.** Predicting peak pinch strength: Artificial neural networks vs. regression. *International Journal of Industrial Ergonomics*, v.18 n. 6, p. 431-441, 1996.
- Favero, L.P.; Belfiore, P.; Silva, F.L.; Chan, B.L.** Análise de dados: modelagem multivariada para tomada de decisões. 2.Ed. Rio de Janeiro: Elsevier, 2009.

Ghiassi M., Nangoy S. A dynamic artificial neural network model for forecasting nonlinear processes. *Computers and Industrial Engineering*, v. 57 n. 1, p. 287-297, 2009.

Haykin, S. *Redes Neurais - Princípios e Práticas*. 2 ed. Porto Alegre: Bookman, 2001.

Kirsten, H.A. Comparação entre os modelos Holt-Winters e Redes Neurais para previsão de séries temporais financeiras – Dissertação de pós-graduação em Engenharia de produção e sistemas – Pontífice Universidade Católica do Paraná – Curitiba/PR, 2009.

Kovacs, Z. L. *Redes Neurais Artificiais: Fundamentos e aplicações*. 4 ed. São Paulo , SP – Livraria da Física Editora, 2004 .

Lustosa, L.; Mesquita, M.A.; Quelhas, O.; Oliveira, R. *Planejamento e Controle da Produção*. 1ed. Rio de Janeiro: Elsevier, 2008. 2ª reimpressão il.

Liu, X.; Wijk, O.; Selin, R.; Edström, J.O. Oxidizing Dephosphorization of Ferro-manganese. *Scandinavian Journal of Metallurgy*, v. 25, n. 5, p. 204-21, 1996.

Lee H.Y., Yang T.H., Chien I.L., Huang H.P. Grade transition using dynamic neural networks for an industrial high-pressure ethylene-vinyl acetate (EVA) copolymerization process. *Computers and Chemical Engineering*, v.33 n.8 , p. 1371-1378, 2009.

Myllykoski P., Larkiola J., Nylander J. Development of prediction model for mechanical properties of batch annealed thin steel strip by using artificial neural network modeling. *Journal of Materials Processing Technology*, v.60, n.4, p. 399-404, 1996.

PedrinI, D. C. ; Caten, C.S. Modelagem estatística para a previsão do teor de fósforo em ligas de ironmanganês. *Ingepro. Inovação, Gestão e Produção.*, v. 2, p. 14-25, 2010.

Perzyk M., Kochanski A.W. Prediction of ductile cast iron quality by artificial neural networks. *Journal of Materials Processing Technology*, v.109 n.3, p. 305-307, 2001.

Santos Filho, E. L. Previsão dos Retornos do Índice Bovespa Usando Redes Neurais Artificiais. 2008. 109p. Dissertação (Mestrado em Engenharia de Produção) – Departamento de Engenharia de Produção e Sistemas, Pontífice Universidade Católica do Paraná, Curitiba, 2008.

Appendix

The regression models proposed by Pedrini and Caten (2010) used for the comparisons:

Table 4 – Regression models

Equation	Model
1	$P = 0,112 + 0,748 P^* - 0,00211 Fe^* + 0,000015 (MnO)^2 - 0,0121 BaO$
2	$P = 0,103 + 0,752 P^* - 0,00212 Fe^* + 0,000017 (MnO)^2$
4	$P = 0,155 + 0,750 P^* - 0,0309 \ln (Fe^*) + 0,000017(MnO)^2$
5	$\ln P = 0,0983 + 0,751 P^* + 0,00000026 (MnO)^3 - 0,000003 (Fe^*)^3 - 0,0128 BaO$
6	$\ln P = 0,501 + 0,163 \ln (P^*) + 0,0000003 (MnO)^3 - 0,000003 (Fe^*)^3 - 0,00381 \ln (BaO)$
8	$\ln (P-P^*) = - 0,804 \ln (Fe^*) + 0,371 \ln (MnO) - 0,656 \ln (CaO)$