ESANN'2008 proceedings, European Symposium on
Artificial Neural Networks - Advances in Computational Intelligence and Learning
Bruges (Belgium), 23-25 April 2008, d-side publi., ISBN 2-930307-08-0.

Multilayer Perceptron to Model the Decarburization Process in Stainless Steel Production

Carlos Spínola[1], Carlos J. Gálvez-Fernández[1], José Muñoz-Pérez[2], José Mª Bonelo[3], Javier Jerrer[4], Julio Vizoso[3]

(1) Department of Electronics,
(2) Department of Languages and Sciences of the Computation,
Málaga University
Edificio Tecnologías de la Información. Campus de Teatinos s/n. 29071 Málaga
Tel: 95-2131388, Fax: 95-2132782, e-mail: cspinola@uma.es
(3) Department of Information Systems,
(4) Melting Section
ACERINOX, S.A.
Fábrica del Campo de Gibraltar, 11370 Los Barrios (Cádiz)
Tel: 956-629305, Fax: 956-629306, jbonelos@acxgroup.com

Abstract: The Argon-Oxygen Decarburization (AOD) is the refining process of stainless steel to get its final chemical composition through several stages, where tons of materials are added and oxygen and inert gas are blown. The decarburization efficiency and the final temperature in each stage are two important values of this process. We present in this paper an empirical model, based on Multilayer Perceptron, to predict these values in order to automate and enhance the production performance of the AOD. Two architectures are proposed and compared.

Keywords: Multilayer Perceptron, AOD, stainless steel, decarburization.

1 Introduction

Advances in Neural Computation have been continuously increasing during the last years and new fields in industrial applications have been developed with this technology, such as the fabrication of stainless steel, a product whose importance has grown after the Second World War, as the evolution in its production volume shows from 1 million metric tons melted by 1950 to 22 millions in 2004 [1].

The chemical refining and decarburization process is a key issue in stainless steel production. This process is performed in the AOD [2] (Argon Oxygen Decarburization) converter (fig. 1), a vessel built to contain up to 150 Tons of liquid steel, where the aimed chemical composition of the steel is achieved, reducing its carbon concentration (decarburization) until 0.01%.

In this paper we describe the design of an empirical model of the decarburization process in the AOD, based on Neural Networks. Its purpose is to predict the oxygen volume that must be injected into the molten metal in the AOD vessel to reduce the carbon concentration until the desired one. The final temperature

ESANN'2008 proceedings, European Symposium on
Artificial Neural Networks - Advances in Computational Intelligence and Learning
Bruges (Belgium), 23-25 April 2008, d-side publi., ISBN 2-930307-08-0.

at the end of this chemical reaction can also be predicted. 2 architectures are proposed and compared.

2 AOD's Decarburization Process description.

The reduction of carbon content is performed in the AOD by a controlled injection of oxygen mixed with inert gas (argon or nitrogen) through submerged tuyeres (fig. 1). This is accomplished in several blowing stages, each one characterized by a constant gas flow and a predetermined oxygen/inert gas volume ratio. In this way, the molten metal is decarburizated with minimum unwanted metallic oxidation [1].

The oxidation reactions are exothermic and produce an important temperature increase, from the initial 1.400-1.500 ºC to more than 1.700ºC. Since this high temperature can damage the refractory internal lining of the AOD shell, cooling scrap is added, making good use of the extra thermal energy obtained. Along this process some material are also added to get the target exact chemical composition of the metal regarding Si, Cr, Ni, Mn, Mo, Cu, etc., elements that, in addition to Fe, compose the different types of stainless steel. Also limes are added to control the slag viscosity.

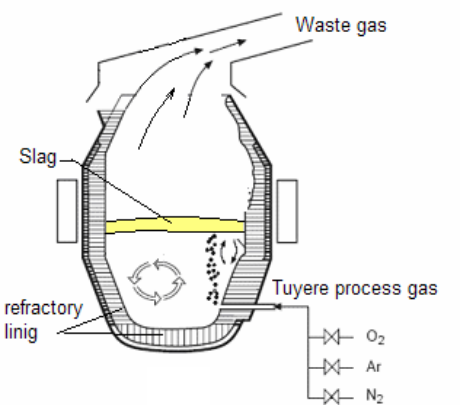

Figure 1: Scheme of the AOD.

All these operations complicate the building of a theoretical analytical model which could fully describe the coupled thermo-metallurgical/fluid dynamic reactions that take place inside the AOD, although important efforts have been made to construct it [3-5]. But they are of limited practical application to our converters as the parameters that these models depend on are difficult to know or estimate in industrial applications [6].

The decarburization efficiency, d_{ef}, is defined as the percentage of the total oxygen blown that reacts with C and Si. The remaining oxygen is combined with Fe, Cr, Mn, etc, forming oxides, which should be reduced in a subsequent phase to recover the metals. If d_{ef} is predicted, the total amount of oxygen needed to reduce the %C can be calculated.

The final temperature, T_f, must be also known in order to decrease it by adding cooling scrap if necessary. So d_{ef} and T_f are the variables that should be estimated by a decarburization model in an industrial application.

3 Building the Empirical Model of the decarburization process

In order to determine d_{ef} and T_f an empirical model of the AOD converter was built by the Acerinox factory melting team. For this purpose a data population was gathered with information about the production process: the initial conditions of the molten steel (chemical composition, initial temperature, mass, etc), the performed

operations (raw materials, lime, and scrap added, oxygen and inert gas flow applied, etc) and the final conditions (%C and the final temperature). The temperature was directly measured in the bath by means of a pyrometer cane at the same moment in which a sample was taken for the chemical laboratory. This pyrometer cane has a nominal error of 4ºC, which grows considerably with temperatures above 1750ºC.

There are more than 2300 cases, with different steel grades and conditions, collected during several years. Each case tries to represent one stage of the decarburization process for a production unit or "heat". It consists of more than 41 variables, including T_f and d_{ef}.

Firstly, a model based on multidimensional linear regression was built. It was used to refine the metallurgical model, adding and removing different input variables. But now a new model, based in the same population but with Neural Networks technology has been implemented to estimate the d_{ef} and T_f. A comparison of both techniques has been studied in [7] and a general superiority of Neural Networks over classical multiple linear regression was concluded.

An analysis of the independent variables has been done to divide the data population into groups with similar characteristics, so that the estimation method could work better for each one of them. For this purpose Principal Component Analysis (PCA) [8] was used. The application of this method to our data is more convenient since many input variables are correlated.

A detailed study of PCA and the application of k-means algorithms to divide the data are presented in [7]. The population was divided in two groups. Fig. 2 shows a projection of data population on the two principal components obtained.

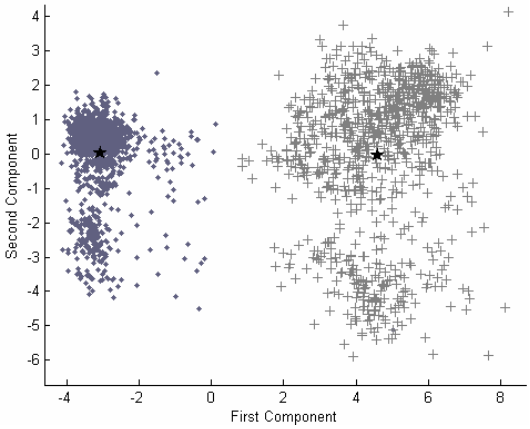

Figure 2: Projection of data population and cluster centres over two principal components.

Thus, the first group contains 951 cases of the first blowing stage, where the ratio oxygen/inert gas is higher, and the second one groups 1440 cases of the following blowing stages.

4 A Perceptron for estimation of average decarburization efficiency and final temperature

4.1. The model

Since the Multilayer Perceptron (MLP) is considered a universal approximator [9], we have designed one MLP to predict d_{ef} and another one to estimate T_f. The MLP architecture (fig. 3) has a set of inputs variables, a hidden layer with an accurate number of neurons and a single neuron in the output layer. Firstly, 39 variables have

been used as inputs and in a second step those ones that are not correlated with the dependent variable have been eliminated, maintaining the training and validation errors.

The hyperbolic tangent transfer function was used for the neurons in the hidden layer and the linear transfer function for the output neuron.

Different numbers of neurons in the hidden layer were used. The back propagation and Levenberg–Marquardt training algorithms [10] were used. The second one provided the best results. To validate the training results of the MLPs and to test its prediction capability, the cross-validation strategy was used. The whole population was randomly divided into two sets: 85% for training and 15% for validation.

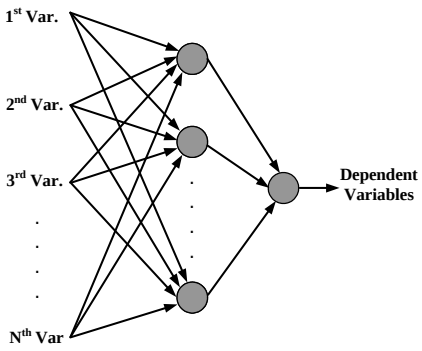

Figure 3: Architecture of the MLP designed with one hidden layer.

4.2. Training and prediction

Since the population was split into two groups, two MLP to estimate d_{ef} and other two to estimate T_f have been designed and trained.

In order to implement each MLP, we choose as input only the variables that can be known at the beginning of a blowing stage. But it must be noted that some information regarding to the aimed final status could also be used as known variables. This is the case of the "desired" final %C at the end of the stage, which can be used as input variable since the operator knows the aim value of carbon concentration which depends on the stage and type of steel to be produced. So, the variables have been divided in the following way:

V_k, V_l : Variables to be predicted (d_{ef} and T_f)

$V_{i \neq (k,l)}$: Variables that can be used as input (known variables at the beginning of a blowing stage).

To model the decarburization process, we have designed two Perceptron, namely:

MLP_E: Perceptron to predict the average decarburization efficiency

MLP_T: Perceptron to predict final Temperature

When there are two or more dependent variables, it is usual to simultaneously predict them all using only $V_{i \neq (k,l)}$ input data (parallel architecture, fig. 4.a). But according to the nature of our problem, we can firstly predict the dependent variable d_{ef} and secondly T_f (sequential architecture, fig 4.b). This is the case because the final reached temperature also physically depends on the average decarburization efficiency obtained in the stage. So we could use the real d_{ef} as input data to train the MLP_T perceptron. But when we use both perceptrons to predict, we should use a sequential approach. We first estimate efficiency d_{ef} (V_k) and then this value can be

used as input variable to the MLP_T to predict the second dependent variable V_l (final temperature). In this way, T_f is always predicted after d_{ef} has been estimated.

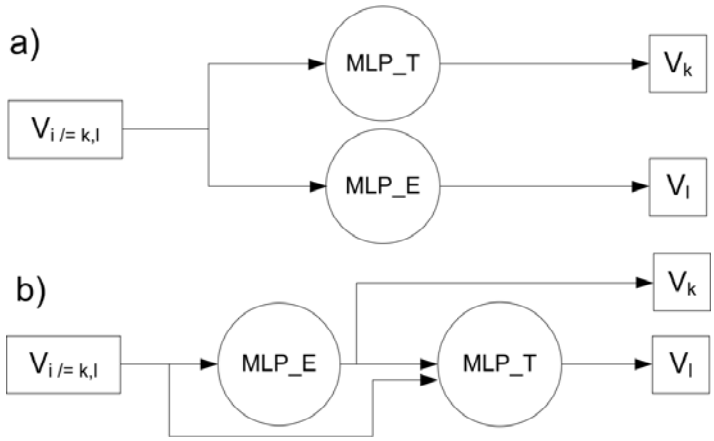

Figure 4: a) Parallel Architecture. b) Sequential Architecture

We have compared the error in temperature prediction with both architectures, i.e. using only $V_{i \neq (k,l)}$ as input variables of both perceptrons (fig. 4.a) and using the calculated V_k as input of the second perceptron (fig 4.b). We have found a reduction in validation error when the sequential architecture is used (see tables 1 and 2). It can be seen that using V_k as input, the average absolute error of the validation data is reduced 1,4ºC in the first group (from 12,3 to 10,9) and up to 2ºC in the second one (9,2 to 7,2). Furthermore, the maximum absolute error has been also reduced in both groups (12ºC and 14ºC respectively).

Group	Input variables		Neurons in the hidden layer	Training data (85%)		Validation data(15%)	
	type	No.		Avg.	Max.	Avg.	Max.
1	$V_{i \neq (k,l)}$	27	6	11,1ºC	49,3ºC	**12,3ºC**	59,1ºC
2	$V_{i \neq (k,l)}$	25	2	8,1ºC	62,5ºC	**9,2ºC**	58,1ºC

Table 1: T_f average and maximum absolute error with the parallel architecture model (using $V_{i \neq (k,l)}$)

Group	Input variables		Neurons in the hidden layer	Training data (85%)		Validation data (15%)	
	type	No.		Avg.	Max.	Avg.	Max.
1	$V_{i \neq l}$	28	4	10,4ºC	46,8ºC	**10,9ºC**	47,2ºC
2	$V_{i \neq l}$	26	2	6,3ºC	53,7ºC	**7,2ºC**	44,2ºC

Table 2: T_f average and maximum absolute error the with the sequential architecture model (using $V_{i \neq l}$)

The average and maximum absolute error for d_{ef} estimation are shown in table 3. Average validation errors of 2.5 and 6.3 for d_{ef} are quite good in this industrial context, where the mean value of efficiency is around 60 percent. Nevertheless there

are several cases with an error well beyond the average and represent atypical cases in which operational problems could have been happened.

Stage	Decarburization efficiency (d_{ef})			Neurons in the hidden layer	Training data (85%)		Validation data (15%)	
	Avg.	Min.	Max.		Avg.	Max.	Avg.	Max.
1	58.4	30,0	89,9	4	1,8	18,4	2,5	12,6
>1	51,2	1,0	118,6	6	5,0	29,0	6,3	32,7

Table 3: Average and maximum absolute error in the estimation of d_{ef}

5 Conclusions

An empirical model has been proposed for decarburization of stainless steel in AOD based on a Multilayer Perceptron to predict the average decarburization efficiency and final temperature.

By using a sequential architecture, the average and maximum error of the predicted final temperature for each production stage is reduced. Although this prediction is fairly good for such a complex process as the stainless steel decarburization, there are some cases with high estimation error, which could be due to errors in the acquisition of data population or operational practices very different from the standard ones.

So, a new data population is now being gathering with more control in order to improve its quality and to enhance the estimation accuracy.

References

[1] Annual Conference of International Stainless Steel Forum in Seoul. Published in "Stainless Steel growth to continue", 17 May 2004, International Iron and Steel Institute, (www.newmaterials.com).

[2] E. T. Turkdogan, R. J. Fruehan, Fundamentals of Iron and Steelmaking, The Making, Shaping and Treating of Steel, 11th Edition. AISE Steel Foundation, 1998.

[3] Ji-He Wei and De-Ping Zhu, Mathematical Modelling of the Argon-Oxygen Decarburization Refining Process of Stainless Steel: Part I y II, Metallurgical and materials transactions B, Vol. 33B, pp.111-119 and 121-127, 2002.

[4] T. DebRoy and D. Robertson, A Mathematical Model For Stainless Steel Making. Part I y II, Ironmaking and Steelmaking, No. 5, pp. 198-206 and 207-210, 1978.

[5] B. Deo and V. Srivastava, Process Control and Optimization of the AOD Process Using Genetic Algorithm, Materials and Manufacturing processes, Vol. 18, No. 3, pp. 401-408, 2003.

[6] G. Staundinger, J.Muller et al., Design Aspects of an Ideal Oxygen Steelmaking Converter, Iron & Steel Technology, Jan 2004, pp.45, ss.

[7] C. Spinola, C.J Gálvez-Fernández, J. Muñoz-Pérez, L. Gálvez-Alcaraz et al, Multilayer Perceptron model to estimate the decarburization efficiency of stainless steel in the AOD process. Engineering Application of Neural Networks International Conference, Lille (Francia), August (2005).

[8] I. T. Jolliffe, Principal Component Analysis (second edition), Springer, 2002.

[9] Barron A.R, Universal approximation bounds for superpositions of a sigmoidal function. IEEE Trans Information Theory, 39, pp 930-945, 1953.

[10] S. M. Burney, T. A. Jilani, A comparison of first and second order training algorithms for Artificial Neural Networks. International journal of Computational Intelligence vol.1, pp.218-224, 2004.