ISIJ International, Vol. 44 (2004), No. 3, pp. 573–580

Blast Furnace Hot Metal Temperature Prediction through Neural Networks-Based Models

Juan JIMÉNEZ, Javier MOCHÓN, Jesús Sainz de AYALA[1)] and Faustino OBESO[1)]

Centro Nacional de Investigaciones Metalúrgicas C/Gregorio del Amo, 8, 28040-Madrid, Spain.
1) Aceralia (Veriña), Aptdo, 570, 33208-Gijón, Spain.

(*Received on June 26, 2003; accepted in final form on November 12, 2003*)

Blast furnace hot metal temperature prediction, by mean of mathematical models, plays an interesting role in blast furnace control, helping plant operators to give a faster and more accurate answer to changes in blast furnace state. In this work, the development of parametric models based on neural networks is shown. Time has been included as an implicit variable to improve consistency. The model has been developed departing from actual plant data supplied by Aceralia from its steel works located in Gijón.

KEY WORDS: ironmaking; blast furnace; neural networks; forecasting; simulation; hot metal temperature.

1. Introduction

The processes that take place inside a blast furnace are highly complex. Some inner processes and inner variables that remain unreachable for blast furnace probes can be mathematically modelled as a mean to achieving at least an approximate knowledge of their behaviour. There are many examples of mathematical models applied to the steel industry in general and to the blast furnace in particular.[1–5)]

Hot metal temperature is, among others, an important parameter to be controlled during blast furnace operation. It should be maintained around a set point that has been previously established. In this way, a stable performance for the blast furnace and a homogeneous quality for the pig iron are achieved.

It would be desirable to develop models able to predict the temporal evolution of hot metal temperature, taking as starting points blast furnace input variables. Several attempts have been made to develop such types of models departing from first principles and numerically solving the thermodynamic equations that describe the process.[1)]

These approaches have provided a great deal of information on the internal blast furnace conditions, and have proved useful for the improvement and stabilisation of its operation. Nevertheless, this approach present a serious difficulty derived from the high complexity of these equations that leads to a numerically unaffordable problem or a solution of such high a cost that it is useless to be applied in real time. One possible solution would be to simplify the equations, but this generally yields poor results.

A possible alternative, especially useful in complex systems is to employ the so-called parametric model. In such an approach the goal is to obtain a function that by means of the estimation of some selected parameters will be able to approximate the relationships among the system inputs and outputs.[6)] Models based on artificial neural networks (ANN) belong to this second kind of model. They have been used extensively in recent years due to their multiple advantages: they are easy to program, present a good adaptation to non-linear systems, their parameters can be calculated on-line (training the ANN), are robust against noise and are easy to reprogram to adapt to changes in the system conditions.[4)]

Along with its advantages, this kind of model, however, presents certain inconveniences: it is necessary to employ a large number of data from the plant that is to be modeled. Otherwise, it is not possible to train properly the ANN. In addition, it is not possible to use the model obtained to extrapolate the behaviour of the system outside the range of values employed to train the network.[4)]

This paper shows the results obtained from different approaches to the problem of hot metal temperature forecasting employing models based on artificial neural networks.

2. Selection of Suitable Variables

In the blast furnace hundreds of process variables are monitored and stored. Not all have a direct influence on hot metal temperature, so it is necessary to perform a prior selection of variables. The set of variables selected as inputs for the models developed in the present work, were some blast parameters along with the ore to coke rate. All of which can be directly controlled by plant operators and have a precise set point assigned. In fact, this set of variables are actually employed to control the blast furnace.

Hot metal temperature, measured in the iron runner at the exit of the tap hole, was employed as the output variable.

Figure 1 shows the values of the variables described for a period of 48 h. The sampling time used was 2 min. As can be seen, blast volume, blast temperature, blast moisture, oxygen enrichment, pulverised coal injection rate and ore to

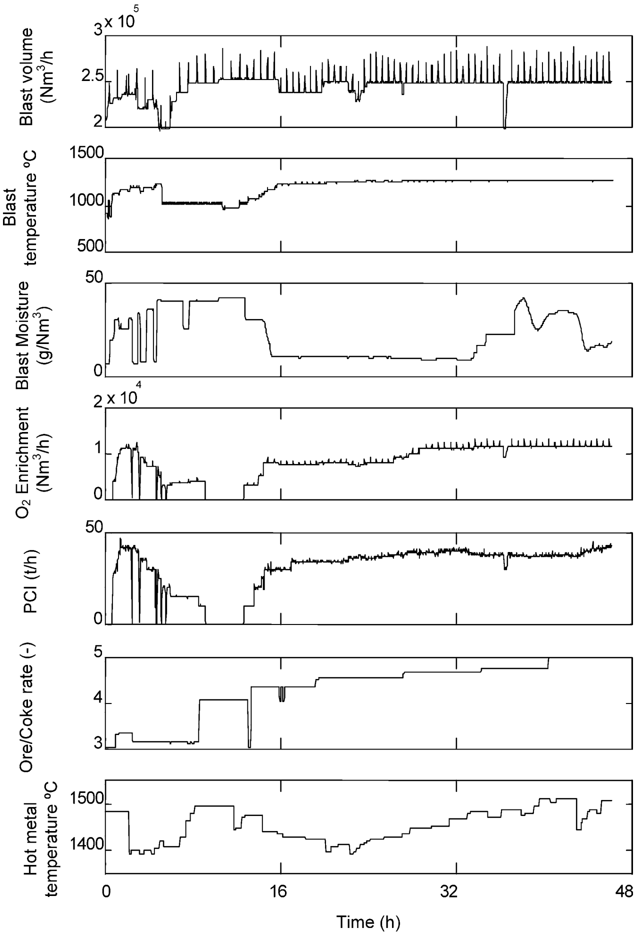

Fig. 1. Variables employed in the model, time evolution during a period of 48 h.

coke rate, was the variables selected as input. The control effort exerted over some of these input variables by means of step changes is clearly observable. For example, the correction applied to blast moisture during the period comprised between the 16 and 32 h it is significant, probably with the objective of restraining the steady decrease of hot metal temperature. Clearly visible as well is the strong correlation between pulverised coal injection (PCI) and oxygen enrichment in the blast. Furthermore, it is possible to check the increase of ore/coke rate in the burden at the same time that PCI increases until it reaches its set point.

Figure 2 shows a more detailed view for hot metal temperature during the same period. In it, the vertical lines and circles represent the beginning of tapping and the time when hot metal temperature was measured, respectively. It is important to emphasise some features of this variable, which are will determine the way in which the problem of its prediction will be addressed.

First of all, it is obvious that neither the tapping rate nor the time when the tapping hole is open is regular. In addition, the number of temperature measurements taken during tapping times and when they are taken are irregular too. Both depend on the plant operator's decision according to their estimation of the state of the blast furnace.

In general, between two and three hot metal temperature measurements are taken by cast. The first is taken shortly after piercing the tap hole, the second when slag starts to flow out and the third near the end of the cast. The process computer records the data when a new measurement is carried out, and repeats this value until the next measurement. This is the reason why the signal obtained for hot metal temperature evolves stepwise.

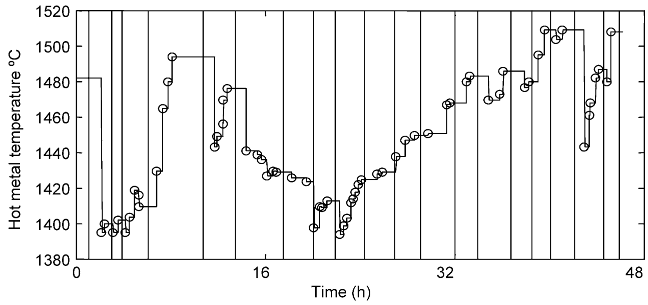

Fig. 2. Hot metal temperature. Circles represent time at which temperature measurements were taken. Vertical lines stand for the beginning of new tap.

These specific features must be taken into account for dynamic modelling purposes. The fact that the value of the variable that is intended to forecast can not be measured with a sampling rate equal to the inputs variables sampling rate, makes it necessary to pre-treat it before being introduced in the model, in order to obtain regularly distributed values. This can be done by means of interpolation among the data available.[7] Another possibility is to include the time explicitly in the model. This latter was the approach employed in this work.

3. Model Development

The structure chosen for the model can be considered as belonging to a class of model known as NARX models (Non-linear AutoRegresive models with eXogenous input). A detailed description of these kinds of models can be found in the literature.[8] The general equation that describe NARX models is as follows:

$$\hat{y}(k)=F(y(k-1),\ y(k-2),\ldots,\ y(k-p),\ u(k),\ u(k-1),\ldots,\ u(k-q)) \quad \text{(1)}$$

The forecasted value of variable y for sampling time k is considered as a function of the values taken by this variable at the p previous sampling times and of the values of the input variables, generically represented by u at the previous sampling times q.

It is clear that such a model must be adapted to take into account the specific system features under study and, in particular, the irregular sampling rate of hot metal temperature measurements. To achieve this, time was explicitly included in the model, according to the following criteria:

Elapsed time between tappings, elapsed time from the previous temperature measurement and elapsed time from the beginning of the tapping until hot metal temperature was measured were included as input variables.

The main reason for including these times was to take into account in some way the residence time of hot metal in the hearth. In general, with the exception of the case where the blast furnace was cooling quickly, hot metal temperature measurement obtained at the beginning of the tapping tends to be lower than the successive measurements. This is due to the fact that the first hot metal comes from the lower

Table 1. Input and output variables for the model. The delay groups were suggested by Aceralia plant operators.

Variable	Units	Suggested Delay groups (τ.) (h.)					
		1	2	3	4	5	6
Hot metal temperature (*T*)	° C	-	-	-	-	-	-
Elapsed time since previous temperature measurement (*tta*)	Min.	-	-	-	-	-	-
Time between tappings (*thc*)	Min.	-	-	-	-	-	-
Elapsed time since the beginning of the tapping (*tic*)	Min.	-	-	-	-	-	-
Blast Volume (C)	Nm^3/h	2	2	2	1	1	1
Blast temperature (Ts)	° C	2	2	2	1	1	1
Blast Moisture (H)	g/Nm^3	2	1	½	1	½	½
Pulverised Coal Injection (PCI)	t/h	2	2	2	1	1	1
Oxygen enrichment (O)	Nm^3/h	2	2	2	1	1	1
Ore/coke rate (m/c)	-	8	8	8	4	4	8

part of the hearth, near the walls, and it is cold. As tapping time progresses, hot metal comes straight from the ore melting at the lower part of the cohesive zone, and it is hotter.

By explicitly including the aforementioned times, the problem of regular sampling rate is avoided. Every value obtained for hot metal temperature is introduced in the model, accompanied by the two characteristics time described above: the time elapsed since the opening of the tap hole and the time the tap hole has remained closed.

The rest of the input variables must also be adapted to achieve the requirements of the model. The original data has been sampled at a two minutes rate. In accordance with the structure of the model described in Eq. (1), it is necessary to take a sufficient number of past data for each one of the variables employed. This leads to an increase in the number of inputs for the model, taking into account more than ever that the blast furnace response is, in general, slow. In addition, it is necessary to fit these inputs to the number of outputs available, *i.e.* the number of measurements of hot metal temperature.

The first problem to elucidate is to determine the number of past data for each variable that must be taken. It is expected that this number be related to the response time of hot metal temperature to changes in the input variables. In the present work, these response times were determined according to the experience of plant operators. There is not sole answer; in fact, plant operators supplied a range of possible time responses for each variable.

The characteristic response time for every variable was included in **Table 1**. Six different response times groups were designed, departing from plant operators' experience who supplied the ranges of the delayed time for each variable.

So, for example, the first group establishes a time response of 2 h for changes in blast volume, blast temperature, moisture, PCI, and oxygen enrichment, and a time response of 8 h for the ore/coke rate.

Coming back to the structure of the model described in Eq. (1), the function F that defines the model structure was built in two steps.

In the first step, input variables time averages were used. These averages were calculated by taking the values of each variable from the time t_a at which hot metal temperature was measured and backwards until a period of time double to the estimated response time (2τ) of each variable has been covered. The values obtained (u_i) for the period are weighted using a Gaussian function with its maximum located at the response time estimated for the variable (τ). The average values obtained with this procedure ($\bar{u}_i$) were employed as inputs for the second step of the model that will be described later. These calculations were performed employing Eq. (2), which summarises the described procedure.

$$\overline{u_i}=\frac{\sum_{t=t_a-2\cdot\tau}^{t_a} e^{-2\cdot([\tau-(t_a-t)]/\tau)^2}\cdot u_i(t)}{\sum_{t=t_a-2\cdot\tau}^{t_a} e^{-2\cdot([\tau-(t_a-t)]/\tau)^2}} \quad \text{..............(2)}$$

Every symbol of Eq. (2) has been already described except t, that represents the time at which u_i was sampled.

Equation (2) has two main effects on the original input data: It smoothes the input signal and it shifts forward the original input signal. Both effects can be observed in **Fig. 3**.

It is possible now to rewrite Eq. (1) according to the specific requirements of the model just described:

$$\hat{T}(k)=F(T(k-tta),\ tta(k),\ thc(k),\ tic(k),\ u_i(k),\ u_i(k-1),\dots,\ u_i(k-2\tau)) \quad \text{.................(3)}$$

It is valuable to make some remarks on Eq. (3). First, a single delayed output is included in the model, the last hot metal temperature available, that was measured *tta* seconds before. *tta* is a linear function of k as far as it is the difference between the current sampling time k and the time at which the last measure of T was taken. This variable is set to zero every time that a new temperature is measured. *tic* is also a linear function of k because it represents the difference between the current sampling time and the time at

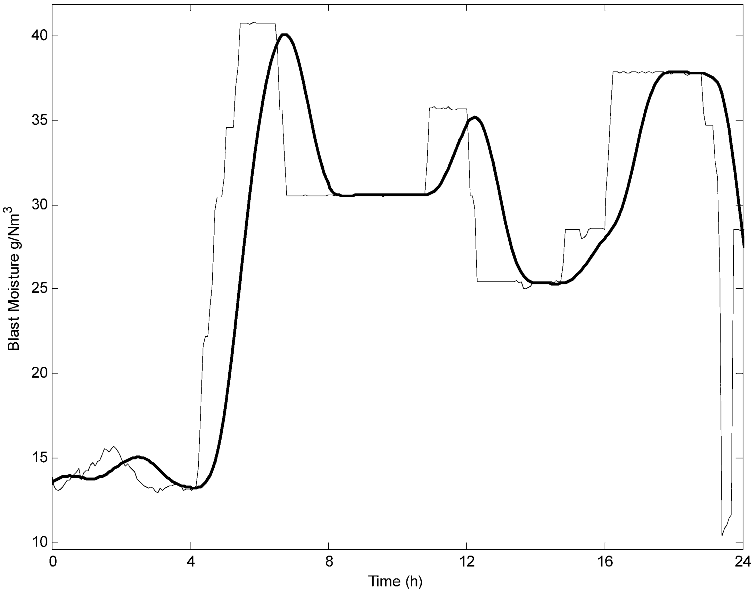

Fig. 3. Comparison between blast moisture data before (thin line) and after (thick line) being averaged by mean of Eq. (2).

which the blast furnace was opened for the last time. It is reset to zero every time that the tapping hole is opened. Finally, *thc* evolves stepwise. It represents the time the blast furnace remained closed between the tapping period at which sampling time k belongs and the previous tapping period. Thus *thc* is reset to zero when the blast furnace is closed, it increases with k until the tapping hole is reopened and it remains constant during the tapping time.

For the rest inputs, generically designated as u_i, the number of delayed inputs depends on each variable, according to the delays suggested in Table 1.

4. Structure and Training of the Neural Network Based Model

For the second step of the function F described in Eq. (1) an artificial neural network was chosen. In the present case, a multilayer perceptron was employed with a single hidden layer.

The multilayer perceptron is an artificial neural network which efficiency as universal function approximator has been widely tested.[9,10]

The input layer has ten neurons. Each one was fed with a different input variable. Six of these inputs were the averaged values obtained in the previous step. The other four values were hot metal temperature taken one step backwards (T), elapsed time between tappings (*thc*) and elapsed time from the beginning of the tapping until hot metal temperature was measured (*tic*) and elapsed time since the previous temperature measurement (*tta*).

The output layer has a single neuron whose numerical output should supply the value of hot metal temperature predicted by the neural network.

A schematic representation of the complete model, including the two steps is shown in **Fig. 4**.

It is interesting to explain at this point that the aim of the first step of the model is to reduce the number of inputs for the neural network. Keeping in mind Eqs. (1) and (3) and the slow blast furnace response, any attempt to implement the model by directly employing neural network directly implies using a large number of inputs. For example, it would be necessary to use 544 inputs for group number one of time responses described in Table 1, considering a sampling rate of 2 min. A neural network with such a high number of inputs requires an excessive computational effort to be trained. Of course, it is possible to employ less delayed samples, but this is equivalent to perform some kind of temporal averaging.

The number of neurons in the hidden layer is one of the variable parameters of the neural network, and must be adjusted according to the dimensionality and complexity of the phenomena to be modelled. In general, as the complexity increased the number of neurons in the hidden layer had to be increased. As far as it is impossible to estimate this feature *a priori*, several trials were performed training neural networks with a different number of neurons in the hidden layer, until an optimum value for this parameter was reached.

Data available were divided into two sets: the first one to train the neural network and the second one to validate its performance. The first set was made up of data from 500 taps and the set employed to validate the results contained data from 150 taps. The training of the neural network was performed employing the Levenberg–Marquard algorithm[10] and the criterion followed in order to stop the training was to obtain a minimum in the normalised sum of squared error (NSSE) for the validation data set.[10]

Neural network with 5, 10, 15, 20 and 40 neurons in the hidden layer were tested for each one of the possible delay time groups described in Table 1. Results for mean square errors obtained at the end of the training have been represented in **Fig. 5**. These results were gathered in response time groups in order to elucidate which set of delays supplied the best approximation.

According to the results showed in Fig. 5, it can be deduced that the better results are obtained for delay group 5. The reason is that for this particular case, the NSSEs obtained for training were similar to the NSSEs obtained for

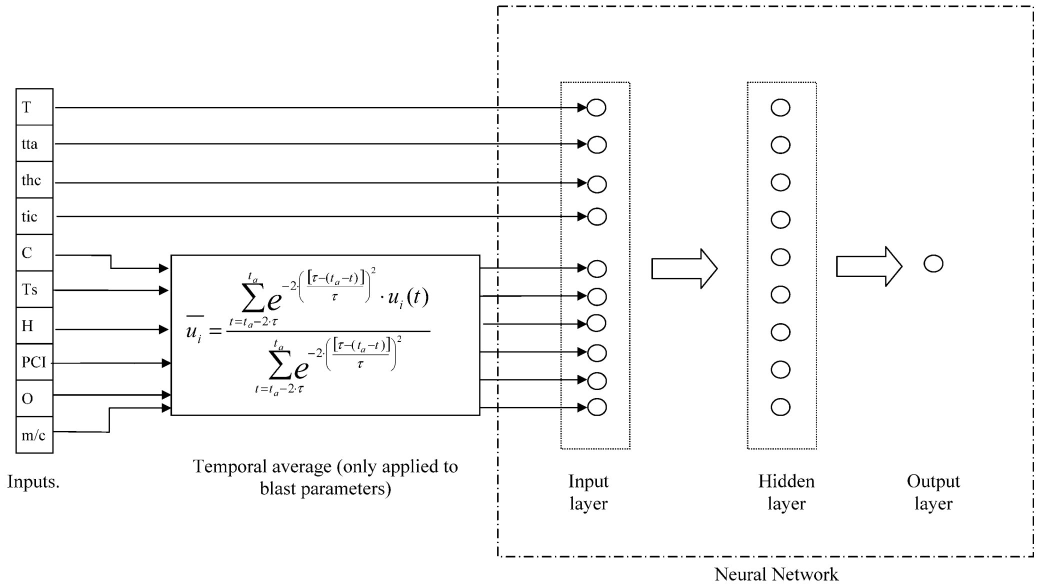

Fig. 4. Model Structure.

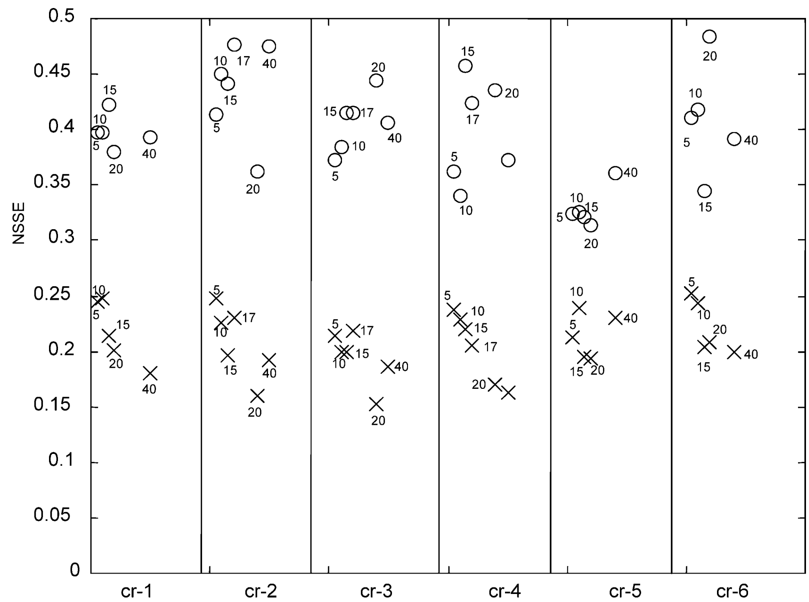

Fig. 5. Normalised sum squared errors. (×) Training errors. (○) Validation errors. The number indicates neurones employed in the hiding layer. Different cases are represented in the horizontal axis.

validation and these latter shows the lowest value.

Figure 5 shows some interesting features of the modelling process. The NSSE obtained for the data set used to train the neural network tended to decrease with the number of neurons employed in the hidden layer. But errors obtained for the validation set do not show the same trend. The reason for this is that when the number of neurons in the hidden layer is increased it becomes easier for the neural network to be over-trained. This phenomenon is related to the fact that the model has more parameters and, as a consequence, more degrees of freedom. So, the parameters are set during training to work with the outputs presented to the model, but these lead to a loss of the generalisation capabilities for the neural network.[10)]

Once case 5 was selected, new neural networks were trained, with different numbers of neurones in the hidden layer, to find the structure for neural network that minimised the errors in hot metal temperature forecast.

Figure 6 shows the results of this second training sequence. As can be seen, the NSSE error reached a minimum for the validation set employing 32 neurons in the hidden layer.

The evolution of the NSSE during training for this network has been represented in **Fig. 7**. The total training period covers 2 500 epochs. One can see that the error obtained for the training data set is greater than the error for the validation data set. This is probably due to the random initial values assigned to the neural network weights. The training was stopped according to the criterion that a minimum NSSE for the validation set was reached. The errors for training and validation data sets are little enough and they are near enough to be considered valid results.

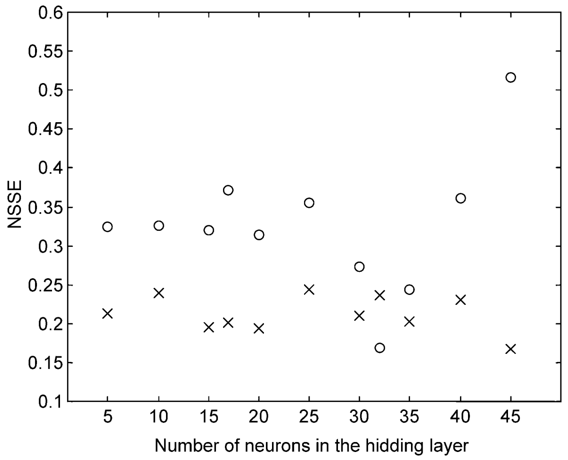

Fig. 6. Normalised sum squared errors for delay group 5. (×) Training errors. (○) Validation errors.

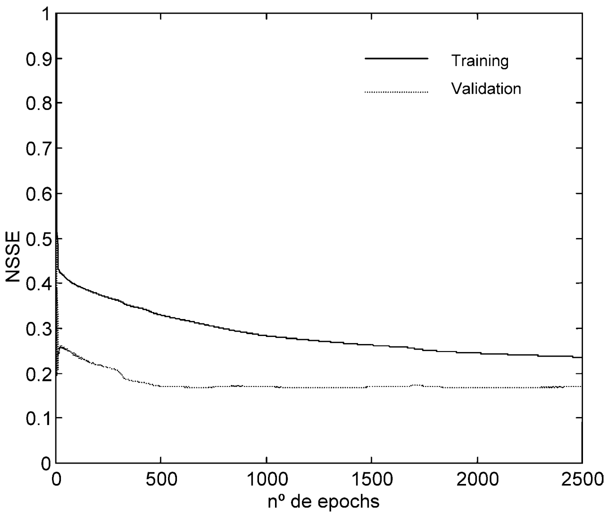

Fig. 7. Normalised errors evolution during a training period of 2 500 epochs. (Delay group 5. Neural network with 35 neurons in the hidden layer.)

5. Results Obtained

Once the neural network structure had been established and the network had been trained the input variable sets were employed to compare their behaviour with the blast furnace performance.

First the neural network was employed to establish a NARX model such as the one described in Eq. (1). This model should be able to know the value of the next hot metal temperature value, departing from the input variables and the last hot metal temperature measured at the blast furnace. Also, a second more suggestive model was established; in it, the last hot metal temperature measured was substituted in the model by its own previous output. The advantages of this kind of model, when possible, are enormous. The model ceases to depend on the previous system output and only depends on its own previous prediction. In this way, the prediction horizon increases, allowing the use of the model to simulate the system.[8)]

This second kind of model is known as NOE (Non-linear Output Error) and its general structure can be represented by Eq. (4):

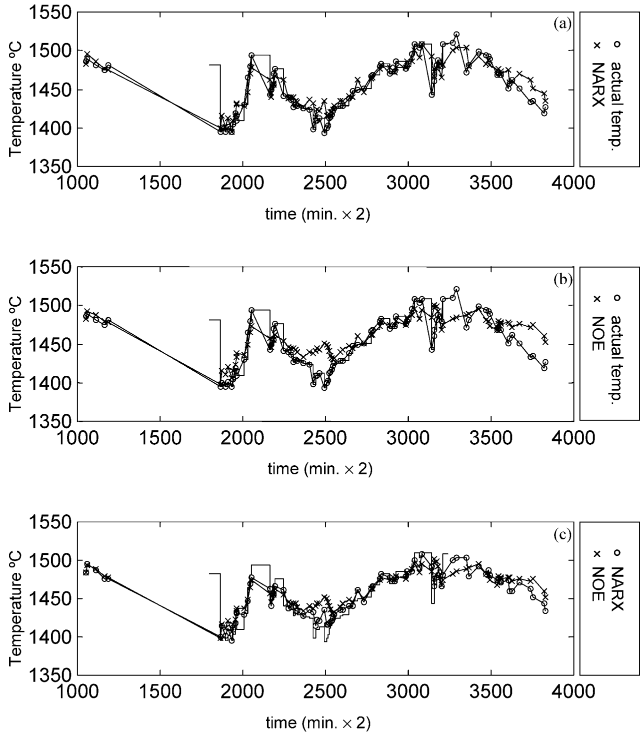

Fig. 8. Results obtained with the models for hot metal temperature predictions. (a) Comparison between actual temperature and prediction from NARX model. (b) Comparison between actual temperature and prediction from NOE model. (c) Comparison between outputs of both NARX and NOE models.

$$\hat{y}(k)=F(\hat{y}(k-1),\ \hat{y}(k-2),\ldots,\ \hat{y}(k-p),\ u(k),\ u(k-1),\ldots,\ u(k-q)) \quad \text{.....................(4)}$$

Figure 8 shows the results obtained with both models for hot metal temperature during a period of 100 h. It should be noted that the data used belongs to the set employed to validate the neural network.

In the period showed in Fig. 8, the first three hot metal temperatures belong to a casting period just before a blast furnace stoppage. Around point 2 000, the blast furnace starts again. The period showed coincides from this point up to 4 000 with the period showed in Fig. 2. The graph of this figure has been superimposed on the result obtained from the model to be compared. The blast furnace undergoes a fast warming followed by a cooling period, reaching a minimum for hot metal temperature around the point 2 500.

Results for both models are comparable and reasonably good. Obviously the NARX model obtains a more precise approximation. Both models have forecast the stoppage correctly, sensibly decreasing the predicted hot metal temperature and following the fast warming of the blast furnace well after the stoppage and the subsequent decrease of temperature.

It is valuable to point out that in the case of the NOE forecast, the whole period was estimated supplying only a single hot metal temperature at the beginning of the period and employing the network's own forecast to obtain the rest. The stability of the response, the simple fact that the model did not oscillate, should be considered as an indication that the model contains at least the basic elements of the system

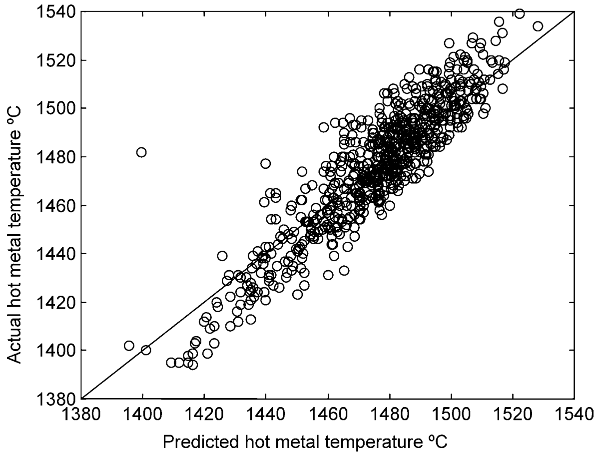

Fig. 9. Comparison among temperature predicted by neural network (NOE) and actual temperature.

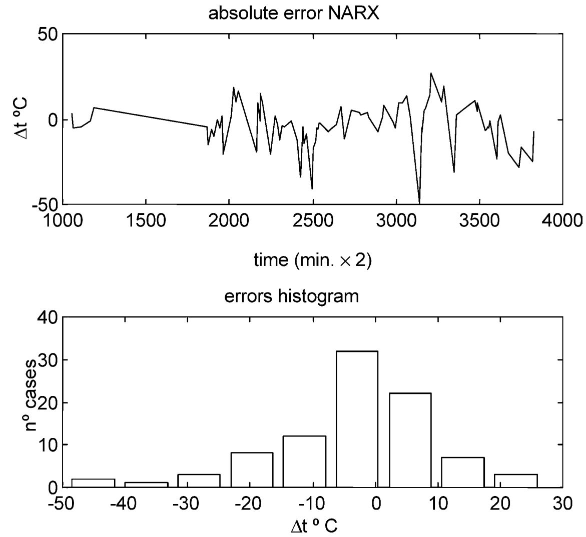

Fig. 10. Prediction error for NARX model.

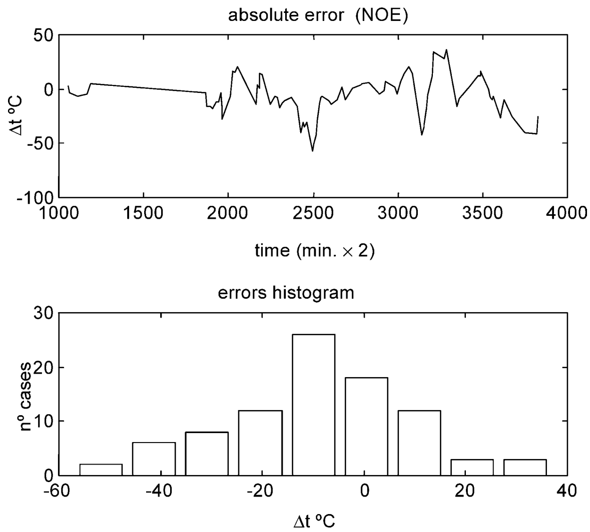

Fig. 11. Prediction error for NOE model.

dynamics modelled.

To estimate the value of the approach the actual value of hot metal temperature has been represented *versus* the predicted values obtained with the NOE model. The graphs includes both training and validation data (**Fig. 9**). The ideal behaviour would be that the points be aligned with the diagonal of the first quadrant. As can be seen, the neural network has a tendency to overestimate the value of hot metal temperature for high temperatures and underestimate it for low temperatures.

Analysis of the prediction errors for both models (NARX and NOE) has been represented in **Figs. 10** and **11**. For a proper analysis of these graphs it is valuable to notice that the mean value of hot metal temperature for the period is 1 500°C, and the variation of temperature is roughly 200°C (between 1 350 and 1 550°C).

For the NARX model, the error histogram shows a deviation of roughly ±10°C. This result is not far from the precision of the sensing device employed to measure hot metal temperature. For the NOE model the range of error is greater, ±20°C (Fig. 11). This result, although worse than the in previous model, can be considered rather good; keeping in mind that the prediction horizon is completely open.

It was pointed out in the introduction, that one of the shortcomings this kind of models is their poor extrapolation ability. Extrapolation may be considered as extrapolation out of the range of input variables employed during training, which includes major changes in the range of variation of input variables employed in the model. Extrapolation may be consider also as extrapolation for a period of time far away from the time at which data were collected to train the neural network. This includes both, the wear of the blast furnace itself and major changes in variables and/or internal processes that were not explicitly included in the model, such as ore and coke properties, heat transfer phenomena, chemical reactions, *etc.* Blast furnace's wear prevents the process from being considered as time invariant for long periods of time. The blast furnace resembles itself only partially as months of campaign go by. Concerning the other not explicitly included variables and internal process; they usually have a slower dynamic and can be considered stable during normal blast furnace operation.

This shows other obvious limitation of the model; it can not supply valid forecast for periods of blast furnace malfunctions, irregular descent of burden, *etc.*

To overcome some of these shortcomings of the model, one possible solution consists on periodically re-train the neural network.

There are different ways of accomplish this. A) re-train the neural network every time that any input exceeds the range previously employed to train the network. B) re-train the network when it begins to cast poor predictions. C) Employ two twin neural networks. One of them is training on line; every time that hot metal temperature is measured, the new data are included in the training data set, the oldest data available are refused and the neural network is re-trained. The updated weights, obtained after re-train, are passed to the second network, which is used only to forecast purposes.

6. Conclusions

The models developed are able to estimate hot metal temperature and its temporal evolution as a function of the input variables that are mainly blast variables.

The viability of neural networks as a dynamic modelling system to forecast hot metal temperature has been shown.

Especially promising is the NOE model, where the model is independent of the previous actual outputs, opening the door to the process simulation. The NARX model should be suitable to develop control systems for the blast furnace.

There are many possible ways to develop models that be able to predict hot metal temperature. The structure suggested is just one of them. The field is wide open and very promising and it is hoped that this work can help other researchers to develop their own ideas in order to reach a better understanding of blast furnace behaviour.

Acknowledgements

The authors thank to the ECSC and CICYT for their financial support and to ACERALIA Company for both the data supplied and their collaboration in this work.

REFERENCES

1) Blast Furnace Phenomena and Modelling, ed by ISIJ, Elsevier Applied Science, New York, (1987), 97.
2) Y. Otsuka, M. Konishi, K. Hanaoka and T. Maki: *ISIJ Int.*, **39** (1999), 1047.
3) M. Falzetti, J. Mochón and S. Kumar: Proc. Application of Artificial Neural Networks Systems in the Steel Industry, ECSC Workshop, European Commission, Brussels, (1998), 9.
4) J. Morris and E. B. Martin: Proc. Application of Artificial Neural Networks Systems in the Steel Industry, ECSC Workshop, European Commission, Brussels, (1998), 9.
5) M. Falzzetti, U. Chiarotti, J. Jiménez, J. Mochón and A. Formoso: Investigation of Fuzzy Logic and Neural Network Based Strategies for Control Optimisation of Ironmaking Process, ECSC Contract N° 7210-AA/421-940-140-422., EUR 19348 en. ISBN 92-828-9221-2. (2000), 24.
6) L. Ljung: System Identification, Prentice Hall, New Jersey, (1999), 140.
7) C. Cantera, J. Jiménez, I. Varela and A. Formoso: *Rev. Metal. Madrid*, **38** (2002), 243.
8) A. Michael, D. E. Henson and Seborg: Nonlinear Process Control, Prentice Hall, New Jersey, (1997), 26.
9) J. Jang, C. Sun and E. Mizutani: Neuro-Fuzzy and Soft Computing, Prentice Hall, New Jersey, (1997), 233.
10) M. Hagan, H. Demuth and M. Beale: Neural Network Design, PWS Publishing Company, Boston, (1995), 11-4, 11-5.