Light Metals 2013 *Edited by: Barry Sadler*
TMS (The Minerals, Metals & Materials Society), 2013

A MIMO Modeling Strategy for Bath Chemistry

Fabio M. Soares[1], Roberto C. L. Oliveira[1]
[1]UFPA – Federal University of Para; Augusto Correa, 1, Belem, PA, 66075-000, Brazil

Keywords: Aluminium Smelting, Process Modeling, Bath Chemistry

Abstract

The bath chemistry is very complex for analytical modeling due to its sensitivity to disturbances from other processes. The control of the electrolyte also represents a great challenge, provided that any change in the heat and mass balance may affect current efficiency and pot life. In order to test a good control strategy, a process model is very helpful. The model allows process experts to design and test process control strategies without compromising reduction cells. In addition a model allows simulations on the process for a long time. In this work we developed a neural-based strategy to model bath chemistry variables in a Multiple Input Multiple Output (MIMO) approach using a data-driven design. The achieved results are very acceptable for the process engineering staff of the industry where this work was performed.

Introduction

Aluminium became essential for the daily life, since many goods and services are available through the application of this light metal. However aluminium production in large scale became viable only after Hall and Heroult had discovered the smelting process which is named after them. It has been evolving since then, allowing for a greater and efficient production. This process is very complex, since it deals with chemical, thermal, electromagnetic and mechanical variables in a set of subsystems highly coupled with each other. This represents a great challenge for this process control, since any change or bias in a subprocess may affect other process. In spite of that, there are many strategies that permit an automatic control in most of cases, such as alumina feeding and anode-cathode distance control. Nevertheless in many smelters there are still many manual activities using measurement equipment subjected to noise and failure due to operations in harsh environments. These factors affect and prevent an efficient control [1].

Under these conditions, models have been helping process teams to better understand, design and test control strategies, prior to launching them into production. A model is a set of mathematical equations on a set of independent (or input) process variables which produce a set of target (or output) process variables. This set of mathematical equations is often called a nonlinear system dynamics model, provided that almost all systems are dynamic and nonlinear. Since the model should work as a virtual plant, it offers the possibility to simulate the process and control strategies. Recent works have shown that neural based models have given good results in nonlinear process modeling. As for aluminium smelting, similar works have been developed using artificial neural networks. Branco [2] developed a model of the electrolytic resistance, a critical control parameter, allowing the test of new feeding strategies and control. Frost and Karri [3] developed models for aluminium fluoride estimation using a number of architectures of neural networks and compared their performance. Soares [4] developed a model for bath temperature following a modeling methodology [5]. Souza [6] developed a method for fluoridated alumina, a not easily measured parameter that strongly affects bath chemistry control.

However, many of these works use an MISO (Multiple Input, Single Output) approach, i.e. they are aimed at inferring only one variable based on a set of variables. In this work we exploit the use of MIMO (Multiple Input, Multiple Output) approach for modeling bath chemistry variables. In its macrostructure, this work is based on Soares [4] and Fontes [7] and presents the following improvements:

- Bath Temperature and Fluoride inferred at the same sample rate;
- Noise and Outlier Filtering applied for cells
- MIMO structure that infers heat balance (bath temperature) and mass balance (aluminium fluoride, calcium fluoride)

This work is structured as follows. In the Introduction the main subject is outlined. Then the Bath Chemistry Challenges are presented, supporting this work's motivation. The concept of Nonlinear models and their techniques is introduced. Subsequently, the methodology used for building the model is presented, including data acquisition, model's parameters estimation and its results. At the Conclusion, final considerations and future works suggestion are provided.

Bath Chemistry Process

Bath chemistry control is really important in the aluminium smelting process, because it strongly impacts on current efficiency [8]. To keep bath composition stable at a good range, some additives are added to the cell, e.g. aluminium fluoride (AlF3 or ALF) and calcium fluoride (CaF2 or CAF). Aluminium fluoride, besides having effect on mass balance, it also affects thermal balance, since it lowers the bath melting temperature and decreases alumina solubility [9].

Bath chemistry control uses aluminum fluoride and bath temperature among other factors to determine how much of any additive should be added into the bath. Although there are many chemical and physical equations to that end, many manual interventions are still needed [10].

An artificial intelligence technique approach to estimate bath chemistry variable arises from the fact that the bath chemistry variables measurements are not available in real, and most of the analytical modeling on both heat and mass balance are inefficient. So, one might consider these techniques to estimate these variables in a shorter basis. In this work we exploit the use of neural networks for bath chemistry modeling on these variables: Bath Temperature, Aluminium Fluoride and Calcium Fluoride.

Nonlinear Models

A real process has variables that represent their dynamic behavior through nonlinear equations. To better understand nonlinear and multivariable models, it is easier first to look at SISO (single input, single output) linear models and their parameters. In its

general form, a linear model consists of a dependent variable y, independent or control variable u and noise or disturbance e, as shown in equation:

$$A(z^{-1})y(t) = z^{-d}\frac{B(z^{-1})}{F(z^{-1})}u(t) + \frac{C(z^{-1})}{D(z^{-1})}e(t) \qquad (1)$$

where *A(.)*, *B(.)*, *C(.)*, *D(.)*, *F(.)* are polynomials of z^{-1}, and z^{-1} is a complex variable whose meaning is also the delay operator [5].
This SISO linear model can be extended to a MIMO form when dealing with vectors *Y(t)*, *U(t)* and *E(t)* instead of single values *y(t)*, *u(t)* and *e(t)*, and matrices instead of the polynomials. For simplicity, let's take into account only the *A(.)* and *B(.)* matrices, then the model with r inputs and p outputs can be rewritten as:

$$Y(t) = A_1.Y(t-1) + \cdots + A_{n_y}.Y(t-n_y) + B_1.U(t-1) + \cdots + B_{n_u}.U(t-n_u) + E(t) \qquad (2)$$

where $A_i \in \mathbb{R}^{p\,x\,p}$ and $B_i \in \mathbb{R}^{p\,x\,r}$, n_y and n_u correspond to the number of delays at the outputs and the inputs respectively. The system variables, now defined as vectors, have the following form:

$$Y(t) = [y_1(t)\ y_2(t)\ \ldots\ y_p(t)]^T$$
$$U(t) = [u_1(t)\ u_2(t)\ \ldots\ u_r(t)]^T$$
$$E(t) = [e_1(t)\ e_2(t)\ \ldots\ e_p(t)]^T \qquad (3)$$

We may also represent all the matrices from the equation (2) in the polynomial form as in equation (1):

$$A(z^{-1}) = I - A_1z^{-1} - A_2z^{-2} - \cdots - A_{n_y}z^{-n_y}$$
$$B(z^{-1}) = B_1z^{-1} + B_2z^{-2} + \cdots + B_{n_u}z^{-n_u} \qquad (4)$$

where I stands for the identity matrix or rank n_y.
The model shown in equation (2) is also known as ARX (AutoRegressive with eXogenous inputs). In practical applications, this model can be rewritten using matrix algebra as:

$$Y(t) = \theta^T\vartheta(t-1) + E(t) \qquad (5)$$

where

$$\vartheta(t-1) = [Y^T(t-1)\ldots Y^T(t-n_y)U^T(t-1)\ldots U^T(t-n_u)]^T$$
$$\theta = [A_1\ A_2\ \ldots A_{n_y}\ B_1\ B_2\ \ldots B_{n_u}]^T \qquad (6)$$

The θ matrix is often called the parameters matrix, while the $\vartheta(t-1)$ matrix is called the regression matrix at the time *t-1*. As can be seen from equations 1-6, the regression matrix is represented by the current known data until instant *t-1*. There are a number of algorithms aimed at finding the optimal parameters of the parameters matrix [5] [11] through several methods, ranging from matrix algebra to gradient descent and other computations.
As for nonlinear counterparts, the model itself is represented by a nonlinear matrix function *F(u(t); θ)*. It is common to use notation of blocks in nonlinear models:

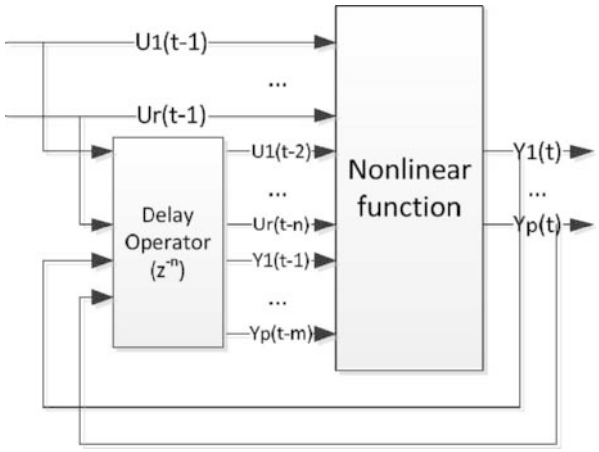

Figure 1. System Block for a Nonlinear Model

The model shown in figure 1 is a model with *r* inputs or independent variables and p outputs or dependent variables, along with *n* input delays and *m* output delays. No noise is being considered for simplicity purposes. The nonlinear function can theoretically be any nonlinear function or structure. The most widely used structure for nonlinear process is the artificial neural network [5][12]. A scheme of an ANN is shown in figure 2:

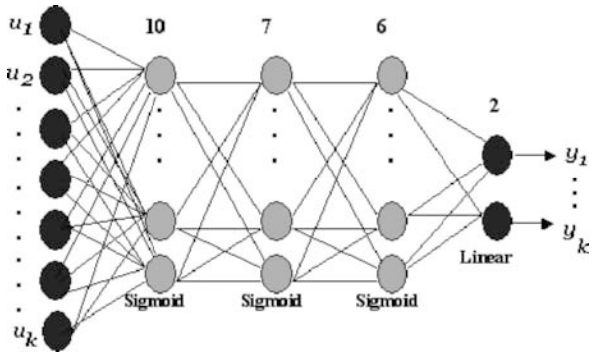

Figure 2. Artificial Neural Network Scheme

In an ANN, all input values are processed by simple units called neurons or neurodes, which are laid out in layers. There are also many architectures of ANN's, but in this work let's take into account only the MLP (Multilayer Perceptron), since many MLP applications in the chemical field are successful [5]. In this architecture, each neuron is connected to all of the neurons of the following layer, and so forth until the output layer which produces the predicted output values. In this context, the connections among neurons represent the model parameters, or weights. Likewise, by changing the neural weights, the output values change as well. Each neuron triggers an output value to the next layer based on an activation function, which can be also nonlinear. This is how the neural network can be used to fit any nonlinear behavior.
The procedure in which the neural weights are adapted is called training, and is performed by a training algorithm. During its execution, the model's predicted values are evaluated against the actual values taken from the plant. The training goes on until

some criteria are met, either by number of iterations (or epochs) or by minimum error.

Model Building

In order to find and build a good model that represents the system's dynamics, some identification procedures described by Fortuna et al [5] and Ljung [11] are followed. A review in the recent literature on similar works in this field is very useful. In this work, the previous works developed by Soares [4] and Fontes [7] have been considered. Summarizing the model building steps, the procedure outline is:

1. **Data Collection and Filtering:** In this stage, good data should be collected and filtered, so that process dynamic may be well represented;
2. **Model structure selection:** from the selected dataset, relevant variables should be selected for the model, besides also eliminating the ones correlated among each other. Also the nonlinear structure should be selected. Commonly, for nonlinear processes, neural networks are chosen;
3. **Model estimation:** In this phase, one should choose how the model parameters will be estimated, i.e. choice of the training algorithm.
4. **Model validation:** An adequate validation strategy must be chosen to verify if the model is responding satisfactorily to new input data.

Data Collection

Aluminium smelters maintain huge database with hundreds of variables. From these, one should determine which data and which variable must be collected. This task requires a deep analysis on the process, and the cooperation with plant experts in the form of meetings [5]. Nevertheless, data statistical analysis such as correlation can also help finding good relevant data for the model, especially when there are a great number of input variables.

In this work we focused on bath chemistry, in which we will produce in a MIMO model values for three variables: Bath Temperature, Aluminium Fluoride, Calcium Fluoride. After reviewing literature and interviewing process experts [1][8][9], we filtered 36 variables that may contribute to the bath chemistry dynamics.

From these variables, we performed linear correlation in order to find variables correlated among each other and those which have a high correlation with the model's output variables. In this work we used the Pearson correlation coefficient, defined in equations (7) and (8):

$$r_{x(k)y(k-\tau)} = \frac{S_{x(k)y(k-\tau)}}{\sqrt{S_{x(k)x(k)}S_{y(k-\tau)y(k-\tau)}}} \quad (7)$$

$$S_{x(k)y(k-\tau)} = \sum_{i=\tau}^{n} x(i)y(i-\tau) - \frac{\sum_{j=0}^{n} x(j)\sum_{j=\tau}^{n} y(j-\tau)}{n} \quad (8)$$

Where

$r_{x(k)y(k-\tau)}$: Pearson correlation coefficient between the variables x and y delayed in τ instants;

$S_{x(k)y(k-\tau)}$: Covariance value between the variables x and y delayed in τ instants;

$S_{x(k)x(k)}$: Covariance value of the variable x.

$S_{y(k-\tau)y(k-\tau)}$: Covariance value of variable y delayed in τ instants;

n: number of samples.

The equation (8) represents the covariance between two variables x and y, taking into account the delay only for one of them. This equation can be modified to determine the covariance between two variables at any delay, as shown in equation (9):

$$S_{x(k-\tau_x)y(k-\tau_y)} = \sum_{i=\tau}^{n} x(i-\tau_x)y(i-\tau_y) - \frac{\sum_{j=\tau_x}^{n} x(j-\tau_x)\sum_{j=\tau_y}^{n} y(j-\tau_y)}{n} \quad (9)$$

If we evaluate the covariance between x and y making τ_x=0, then eq. (9) becomes eq. (8).

After this procedure, we selected the following variables listed in table I.

Table I
Selected Variables for Soft Sensor Model with delays

Variable	Unit	Symbol	Corr. w/ TMP	Corr. w/ ALF	Corr w/ CAF
Alumina Feed	Kg	QAL	-0.677	0.412	0.049
Bath Temperature	℃	TMP	1	-0.88	0.17
Aluminium Fluoride percentage	%wt	ALF	-0.621	1	-0.209
Calcium Fluorite percentage	%wt	CAF	0.21	-0.233	1
Aluminium Fluoride Feed	Kg	ALFA	0.589	-0.522	0.197
Calcium Fluoride Feed	Kg	CAFA	-0.235	0.192	-0.534
Cell Voltage	V	VMR	-0.477	0.44	-0.151
Cell Resistance	μΩ	RMR	-0.52	0.541	-0.19
Metal Level	cm	NME	-0.19	0.23	0.021

We can see through the table that some variables present low correlation with the target variables, and most of them have low correlation values with Calcium Fluoride. However, these variables were kept due to the fact that historically the process team performs statistical analysis using them.

Filtering and Dataset Separation

Filtering is very important to ensure that only good noise-free data are taken for model building. When there are outliers, there is always an indication of malfunctioning in the plant, and they are extremely bad for modeling. In order to find and remove these outliers, we performed the 3 sigma rule filtering, which is suited when large number of outliers is suspected [5]. It is defined by the equation (10):

$$d_i = \frac{x_i - \mu_x}{\sigma_x} \quad (10)$$

Where x_i is the value of x at the *i-th* instant, μ_x is the mean value of x, σ_x is the standard deviation of x, and finally d_i is a weighted distance of the value x at *i-th* instant from the mean. When $|d_i|$ is greater than 3, its corresponding record is considered an outlier. Table II shows each variable's range determined by the 3 sigma rule and percentage of records that fit in that range.

Table II
Selected Variables' Range determined by 3 sigma rule

Variable	Symbol	Lower limit	Upper limit	% of records in
Alumina Feed	QAL	1875.2	2860.2	99.96
Bath Temperature	TMP	932.5	992.6	99.41
Aluminium Fluoride percentage	ALF	4.38	17.73	99.54
Calcium Fluorite percentage	CAF	3.042	6.926	99.91
Aluminium Fluoride Feed	ALFA	-51.54	141.39	99.64
Calcium Fluoride Feed	CAFA	-22.56	34.185	99.69
Cell Voltage	VMR	4.135	4.479	99.13
Cell Resistance	RMR	14.49	16.34	99.27
Metal Level	NME	18.42	24.24	99.87

Data collection spanned over 200 cells from different potlines, totalizing circa 16,000 records. This query was designed to keep full time series from each variable for each cell. That means that any cell presenting at least one outlier in one variable is removed from the dataset. Therefore, applying this rule, we kept only the "good" cells, reducing the dataset down to 6,395 records or only 106 cells.

Then we separated the data for training, tests and validation [12]. The training dataset will be used to estimate the model, the test dataset will test the model during the training procedure, and the validation dataset will evaluate how good the model is predicting values for new data. Table III shows how the dataset was separated into these three subsets:

Table III. Dataset separation.

Training		Tests		Validation	
Cells	Records	Cells	Records	Cells	Records
70	4,212	16	970	20	1,213

Model Structure and Estimation

As can be seen in similar texts in the literature [5][6][7], a good structure for these complex processes' modeling is an NARX (Nonlinear AutoRegressive with eXogenous inputs) structure. In this structure we define how many delays for each variable should be considered for both input and output. This structure should be a nonlinear function to perform the input-output mapping, which in this case was an MLP Neural Network. During the estimation process we set up several three layers neural networks by randomizing its parameters (number of neurons in the hidden layer, activation function, number of delays used). After the training, the ANN with best performance is chosen to represent the model.

We chose the Levenberg-Marquardt [13] as the training algorithm, since it is able to converge MLP neural networks faster than traditional optimization algorithms, such as back propagation and steepest descent. We should highlight that, in order to prevent saturation in the ANN activation functions, we normalized input data according to equation (11):

$$y = \frac{(y_{max}-y_{min})(x-x_{min})}{(x_{max}-x_{min})+y_{min}} \quad (11)$$

where y_{max} is 1; y_{min} is -1; x is the value to be normalized; x_{min} is the minimum value to be normalized; x_{max} is the maximum value to be normalized; y: normalized value.

The parameter used to measure the network performance was the MSE error defined by equation (12):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{P} e_k(i)^2 \quad (12)$$

where N is the size of the dataset, P is the number of the model's output variables, $e_k(i)$ is the error between the predicted output and the actual value for the variable k at the instant i.

Results

After performing the training across over 100 candidate neural networks, we find a good configuration based on the overall error MSE. This configuration is shown in table IV.

Table IV. MLP Configuration for Bath Chemistry MIMO model

Parameter	**Value**
Neurons in 1st Layer	10
1st Layer Act. Func.	Input Layer
Neurons in 2nd Layer	34
2nd Layer Act. Func.	Hyperbolic Tangent
Neurons in 3rd Layer	3
3rd Layer Act. Func	Linear
Delays	3
Training Epochs	6
Normalized MSE	0.0172
MSE for ALF	2.452 %
MSE for TMP	8.39 ºC
MSE for CAF	0.793 %

We can see through both the scatter and time series plots that the model's performance is pretty good:

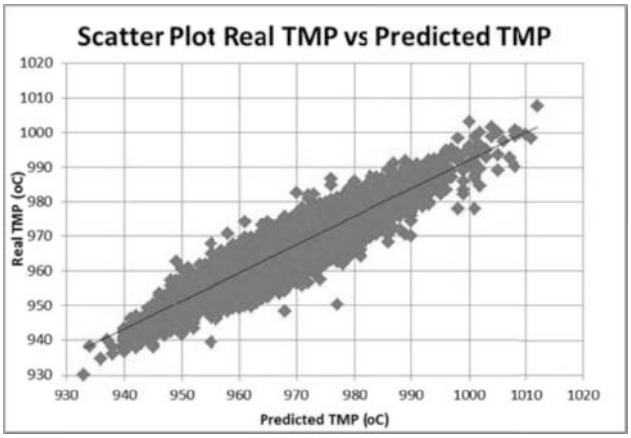

Figure 3a. Scatter plot for Real Bath Temperature and its prediction by the MIMO model

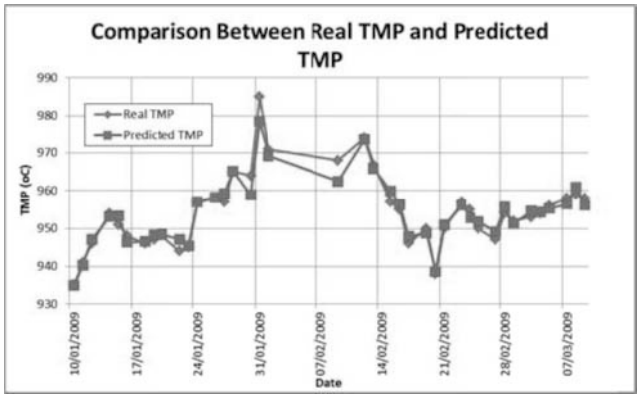

Figure 3b. Time series plot comparing Real Bath Temperature and its prediction by the MIMO model

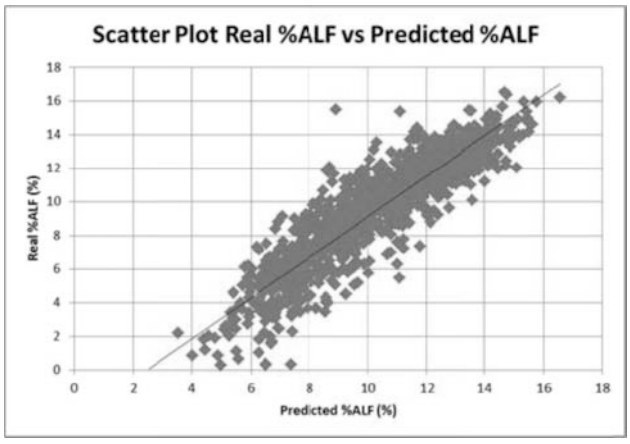

Figure 4a. Scatter plot for Real Aluminium Fluoride and its prediction by the MIMO model

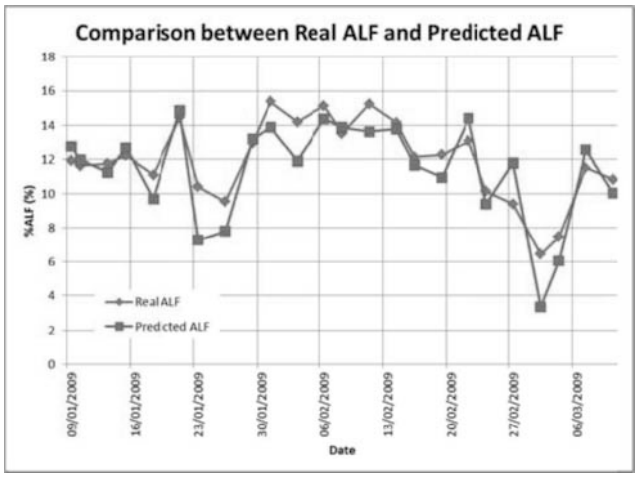

Figure 4b. Time series plot comparison between Real Aluminium Fluoride and its prediction by the MIMO model

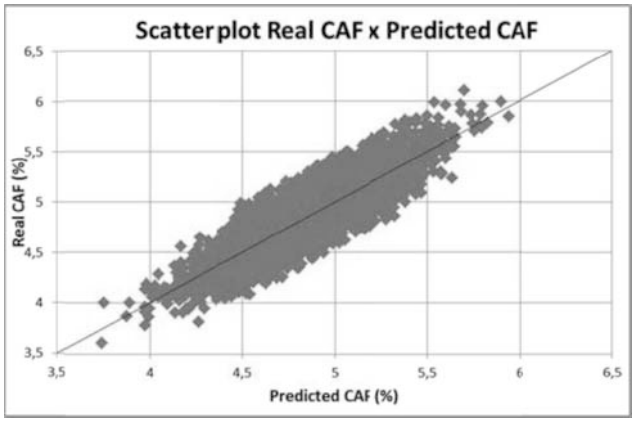

Figure 5a. Scatter plot for Real Calcium Fluorite and its prediction by the MIMO model

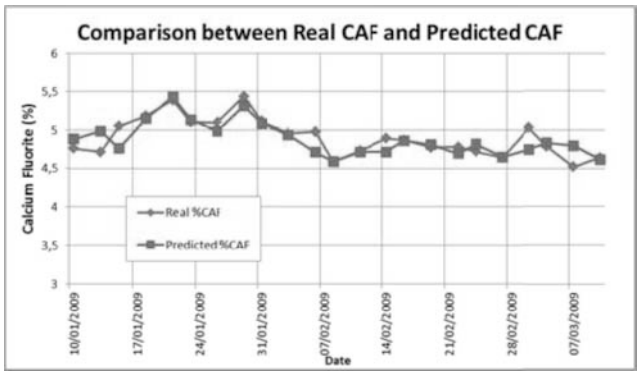

Figure 5b. Time Series plot comparing Real Calcium Fluoride and its prediction by the MIMO model

The works developed by McFadden [9], Hyland [8], Frost and Karri [3], Soares [4], and Pereira [10] have shown the possibility to model bath chemistry by taking into account relevant variables from the cells. In addition, it is known that heat balance is highly related to mass balance [14], so a model to predict these variables at once is absolutely possible. The results just confirm this fact.
With this model, control strategies for both mass and heat balance can be designed and tested across several scenarios. A good application would be to test fluoride additions strategy and evaluating its performance by simulating the cells behavior after a number of steps.

Conclusion

This work proposed Bath Chemistry modeling using a MIMO approach. With this tool, this process can be simulated and control strategies may be tested and its effects may be evaluated prior to launching. The heat balance and mass balance are extremely important variables for guarantee of high current efficiency; nevertheless their measurements require laboratory analysis and/or manual procedures. That motivated the building of a model. The results achieved show that this model can really simulate the cell's bath chemistry behavior. For future works, we suggest the same approach to include more variables from this process and/or the evaluation of control strategies in this model.

Acknowledgements

We acknowledge the company which provided us the data. For confidentiality reasons we are not allowed to disclose the company's name. We also acknowledge the funding from the Brazilian Committee for Research (CNPq) and the Federal University of Pará for the support given to this work.

References

1. L. Grjotheim, and H. Kvande "Introduction to Aluminium Electrolysis. Understanding the Hall-Heroult Process". 2nd Edition *Aluminium* – Verlag Dusseldof (1993), 245-251.

2. M.V.B.M. Branco, V.G. Pereira, R.C.L. Oliveira, M.A.G. Castro, "Neural Model of Electrical Resistance in Reduction Cells of Aluminum to be applied on the Process Control" *TMS Light Metals* (2006), 353-357.

3. F. Frost, V. Karri, "Identifying Significant Parameters for Hall-Heroult Process Using General Regression Neural Networks", *IEA/AIE 2000*, Springer Verlag Berlin Heidelberg (2000), 73-78.

4. F.M. Soares, R.C.L. Oliveira, M.A.G. Castro, "Bath Temperature Inference through Soft Sensors using Neural Networks", TMS Light Metals, Proceedings of the Technical Sessions (2010), 467-472.

5. L. Fortuna, S. Graziani, A. Rizzo, M.G. Xibilia, Soft Sensors For Monitoring And Control Of Industrial Processes , Spinger, 1st edition, (2007).

6. A.M.F. Souza, C.M. Affonso, F.M. Soares, R.C.L. Oliveira, "Soft Sensor for Fluoridated Alumina Inference in Gas Treatment Centers", *IDEAL 2012, Lecture Notes in Computer Science*, vol. 7435, (2012), 294-302.

7. O.Fontes, F.M. Soares, R.L. Oliveira, "Estimation of Aluminium Fluoride Concentration in Aluminium Reduction Cells through a Soft Sensors Approach", *IDEAL 2012, Lecture Notes in Computer Science, vol. 7435 (2012),* 303-311.

8. M.M. Hyland, E.C. Patterson, F.S. McFadden, B.J. Welch, "Aluminium Fluoride Consumption and control in smelting cells". *Scandinavian Journal of Metallurgy* (2002), 404-414.

9. F.J.S. McFadden, B.J. Welch, D. Whitfield, M.S. Kazacos, "Control of Temperature in Alumium Reduction Cells – Challenges in Measurements and Variability". *TMS Light Metals* (2001), 1171-1178.

10. V.G. Pereira, R.C.L. Oliveira, F.M. Soares, "Fuzzy Control Applied to the Aluminium Smelting Process", *Fuzzy Logic – Book 2. Intech online*, Rijeka, Croatia (2012).

11. L. Ljung, "System identification: theory for the user", 2nd edition, T. Kailath, Ed. Englewood Cliffs, Prentice Hall (1999).

12. S.S. Haykin, "Neural Networks: A Comprehensive Foundation" (Hamilton, ON: Prentice Hall, 1998), 43-63

13. M. Hagan, and M. Menhaj, "Training Feedforward Networks With the Marquardt Algorithm," *IEEE Transactions on Neural Networks*, 5 (6) (1994), 989-993.

14. F.J.S. McFadden, G.P. Bearns, P.C. Austin, B.J. Welch,. "Application of Advanced Process Control to Aluminium Reduction Cells". *TMS Light Metals* (2001), 1233-1243.