International Journal of Technology (2015) 6: 480-491
ISSN 2086-9614 © IJTech 2015

ARTIFICIAL NEURAL NETWORK MODELING AND OPTIMIZATION OF HALL-HEROULT PROCESS FOR ALUMINUM PRODUCTION

Sepehr Sadighi[1*], Reza Seif Mohaddecy[1], Yasser Arab Ameri[2*]

[1] *Catalytic Reaction Engineering Department, Catalysis Research Division, Research Institute of Petroleum Industry (RIPI), West Blvd., Azadi Sports Complex , P.O. Box 14665-137, Tehran , Iran*
[2] *Faculties of Engineering, Shahrood branch, Islamic Azad University, Shahrood, Iran*

(Received March 2015 / Revised: June 2015 / Accepted: July 2015)

ABSTRACT

Experience in applying a hybrid artificial neural network (ANN)-genetic algorithm for modeling and optimizing the Hall-Heroult process for aluminum extraction is described in this study. During the stage of modeling, the most important and effective process variables including temperature and cell voltage, metal and bath heights, purity of CaF_2 and Al_2O_3, and bath ratio are chosen as input variables whilst outputs of the model are product purity, ampere efficiency, and product rate. During three years of operation, 19 points were selected for building and training, 7 points for testing, and 7 data points for validating the model. Results show that a feed-forward Artificial Neural Network (ANN) model with 3 neurons in the hidden layer can acceptably simulate the mentioned output variables with the Mean Squared Error (MSE) of 0.002%, 0.108% and 0.407%, respectively. Utilizing the validated model and multi-objective genetic algorithms, aluminum purity and the rate of production are maximized by manipulating decision variables. Results show that setting these decision variables at the optimal values can increase approximately the metal purity, ampere efficiency, and product rate by 0.007%, 0.185%, and 20kg/h, respectively.

Keywords: Aluminum production; Artificial neural network; Hall-Heroult process; Modeling; Optimization

1. INTRODUCTION

Aluminum with a melting point of 2,040°C is a poor electrical conductor and highly reactive. The industrial production of pure aluminum is carried out electrochemically in the Hall-Heroult process, which was discovered in 1889 and it remains to date the only known viable method for the production of aluminum (Prasad, 2000). In this electrolysis process, pure alumina is dissolved in an electrolyte of molten cryolite (Haupin, 1995) in large electrolytic furnaces, called reduction cells (Keniry, 1994). By means of carbon anodes suspended in the electrolyte, the electrical current is passed through the electrolyte mixture causing metallic aluminum to be deposited on the carbon cathode at the bottom of the cell. Carbon anodes are a major part of the cost of primary aluminum production. Therefore, the focus of industrial plants is to minimize the consumption of anode by improving its quality. Hence, the determination of the impact of raw materials, and also process conditions on the baked anode property are so important. Although sufficient data can be collected from these plants, it is complex and difficult to analyze these data using conventional methods such as kinetic-based models, classified as deterministic or first principal models.

* / 2 NJNJ Š ă LJ 2 ƴ Ř λ ƴ Ǝ sadighis@ripi.ir, Tel.: +98-21-48256113, Fax.: +98-21-44739709
Permalink/DOI: http://dx.doi.org/10.14716/ijtech.v6i3.887

Conversely, to model such complex processes, using an Artificial Neural Network (ANN) model, classified as a black box model, this process could be faster and simpler (Sadighi et al., 2013). ANNs are high performance, non-linear analytical tools which are capable of establishing the relationship between the input/output data without prior knowledge (Serra et al., 2003); In some cases, they can even provide better results than empirical correlations (Perazzini et al., 2013). Additionally, ANNs can be applied as hybrid models. The term hybrid modeling is used to describe the incorporation of prior knowledge about the process under consideration in a neural network modeling approach (Bellos et al., 2005; Bhutani et al., 2006).

According to the literature search, research is scarce that aims to develop an ANN model to simulate and optimize the Hall-Heroult process. Frost et al. (2000) proposed a Back Propogation (BP) neural network to predict electrolyte temperatures. In this work, it was shown that careful consideration for training data would give more accurate electrolyte temperatures. In another paper, based on the data collected during 659 days of industrial production, an ANN model for an aluminum factory was presented (Durici et al., 2012). The fitting level obtained using this model resulted in a R^2 value (coefficient of multiple determinations) of 0.723.

The current research aims to propose a method according to the ANN modeling approach to simulate the output variables (i.e., metal purity, production rate, and ampere efficiency), and also optimize the input variables (i.e., temperature and cell voltage, metal and bath heights, and bath ratio) of an industrial scale at a Hall-Heroult plant. This work could be considered significant, due to presenting a practical approach which can be efficiently utilized to simulate and optimize the Hall-Heroult process.

2. PROCESS DESCRIPTION

In 1889, Hall, (USA) and Heroult, (France) invented simultaneously and independently of each other the process to produce aluminum by electrolysis into liquid aluminum in electrolysis cells. In these cells, alumina (Al_2O_3), which is dissolved in a bath of molten cryolite (Na_3AlF_6), is decomposed under the action of a high-intensity (50- 300 kA) and a low-voltage (~4V) DC current. A constant electric current with high-intensity and low-voltage passes through the electrolytic cell from the carbon anode to the bath, and then onto the carbon cathode. The latter is built in the form of a rectangular box to facilitate the gathering of the liquid aluminum as shown in Figure 1.

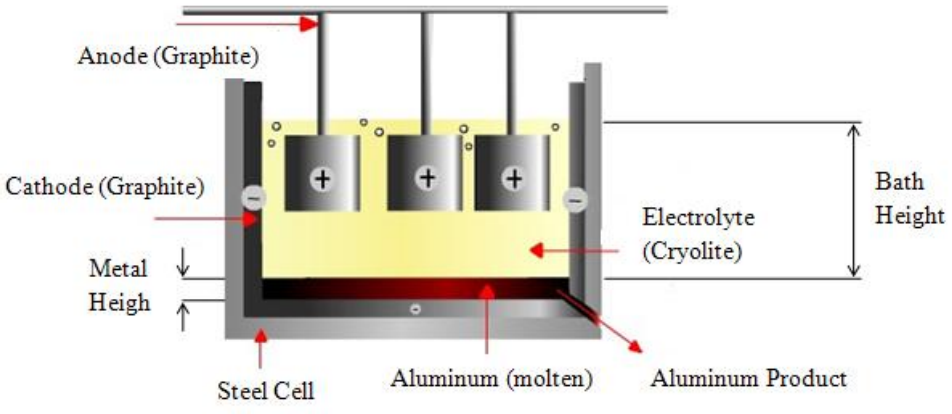

Figure 1 Process flow diagram of the target aluminum factory (Haupin, 1995)

The anode of the electrolytic cell is made of carbon, and the pool of already produced aluminum acts as a cathode. The oxygen of the alumina is discharged at the anode where it reacts with the carbon to produce carbon dioxide (CO_2).

A steel cell which is lined with carbon blocks and thermal insulation material contains the liquid cryolite electrolyte and liquid aluminum. The process uses electrical energy to reduce electrolytically aluminum oxide, and to keep the electrolyte at a temperature of about 950°C. During aluminum production, the chemical reactions continuously consume alumina and anodes, which must be added and replaced, respectively in the electrolytic cell.

Aluminum and anode gases (i.e., carbon dioxide and carbon monoxide) are produced and removed from the cell, respectively. In addition to cryolite, the bath usually contains various additives, mainly aluminum fluoride (AlF_3) and calcium fluoride (CaF_2), with the purpose of improving the physicochemical properties of the bath, and lowering its melting temperature. If the alumina is maintained at too high a concentration, then un-dissolved slurry of alumina may form at the bottom of the cell, which causes the reduction of current efficiency. For adequate operation, the alumina concentration should be maintained in the range of 2- 3%.

3. MODELING METHODOLOGY

One of the most well-known structures for supervised learning of neural networks is the Multi-Layer Perceptron (MLP), which is generally used for classification and prediction problems (Serra et al., 2003). In this method, neurons consist of at least an input layer, an output layer, and one or more hidden layers. For this topology, the information only propagates in the forward direction. The most widely employed networks have only one hidden layer (Hagan et al., 1995). Each node within a given layer is connected to all of nodes of the previous layer. The node sums up the weighted inputs and a bias. Then, it sends results through a linear function as follows (Haykin et al., 1998):

$$a_j = \sum_{i=0}^{m} w_{ji} x_i + b_j \tag{1}$$

where w_{ji} is the weight that goes from the input i to the hidden neuron j ; b is the bias to the node, and x_i is the input of the neuron. By utilizing an activation function (f), the output of the neuron can be written as follows:

$$z_j = f(a_j) \tag{2}$$

This activation function is applied to model the nonlinear process behavior. In this work, the activation function utilized for hidden and output nodes is the tangent sigmoid function as follows:

$$f(a) = \frac{e^{a} - e^{-a}}{e^{a} + e^{-a}} = \tan sig(a) \tag{3}$$

Neural networks have been broadening into numerous types such as radial basis, Kohonen, recurrent, Hopfield, feed forward, neuro-fuzzy networks and so on (Mat Noor et al., 2010). After creating the structure of the ANN, the training procedure is carried out by introducing a set of known inputs and outputs. ANN can learn the trend of these known data by manipulating the weights and biases (Behbahani et al., 2009). These parameters are adjusted using a back-propagation algorithm i.e., iterative reduction in training errors utilizing a generalized delta rule. This procedure is followed until a minimum value for the Mean Square Error (MSE) can be provided. Additionally, the success in obtaining a reliable and robust network depends strongly on the choice of process variables involved in the model (Parisi et al., 2002).

Therefore, it is essential to include all momentous variables, affecting the yield and quality of the product in the input layer of the ANN structure.

In this research, feed forward and recurrent layer networks are used to model the target aluminum extraction plant. Input neurons of the proposed networks are the cell temperature, voltage, purity of CaF_2 and Al_2O_3, metal and bath heights, and bath ratio (see Figure. 2). The latter is defined as the mass ratio of NaF to AlF_3, which are used as additives. The most important output variables included in the output layer are the purity of product, ampere efficiency, and the rate of production.

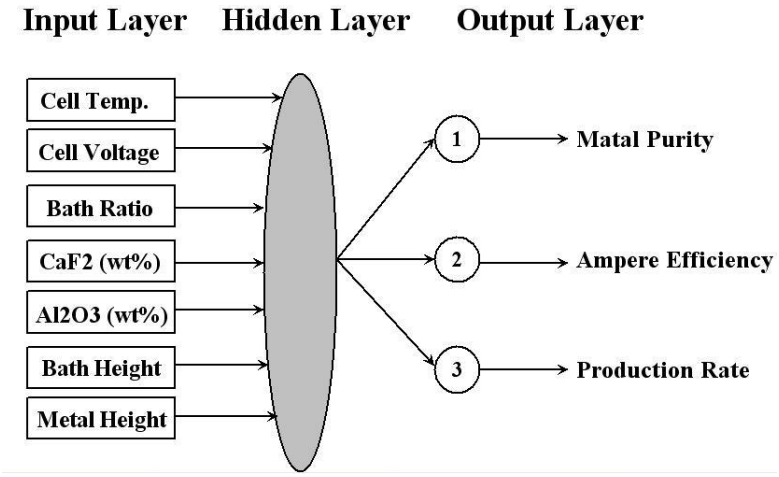

Figure 2 Structure of the ANN model developed for the target aluminum factory

Thirty-three data points were collected from the target plant, and they were divided into three groups i.e., training (60%), testing (20%), and validating (20%) ones. Developed networks are trained, tested and validated using the Neural Network Toolbox (newff and newlrn functions) of MATLAB 2012a. The transfer or activation function used in the hidden and output nodes is the tangent sigmoid function. Training of the ANN is carried out using the 'trainlm' syntax, applying Levenberg-Marquardt back propagation method to estimate weigh

4. OPTIMISATION OF HALL-HEROULT PROCESS

The metal purity and production rate of the plant are identified as significant process variables of the Hall-Heroult process. Therefore, they are maximized simultaneously using the validated ANN model by manipulating the process variables i.e., temperature and cell voltage, metal and bath heights, and bath ratio. CaF_2 and Al_2O_3 purities are dependent to the quality of the feed, and so they are assumed as constant variables during the optimization procedures.

Although there are many numerical methods that have been presented to solve an optimization problem, in this paper the genetic algorithm (GA) has been chosen to optimize the target Hall-Heroult process. GA is a part of soft computing, a branch of computer science that deals with exploring the search space, selecting the best solution, and working for global optimization (Joshi, 2014). These algorithms (GAs) have been applied to a variety of function optimization problems, and were shown to be highly effective in searching large and complex response surfaces even in the presence of difficulties, such as high dimensionality, multimodality and discontinuity (Goldberg, 1989). The GA operates on a population of potential solutions, using the principle of survival of the fittest to produce successively better solutions to a problem. At each generation of a GA, a new set of answers is created by the process of selecting individuals according to their level of fitness in the problem domain and regenerating them using operators

mimicked from natural genetics. This process leads to the evolution of populations of individuals that are better suited to their environment than the individuals from which they were created, just as in natural adaptation.

5. RESULTS AND DISCUSSION

During three year period, under normal operating conditions, actual data were obtained from the target plant. Since consumption of carbon anode was approximately constant versus the time of operation (0.45 per kg of pure aluminum), the cell age was not considered as an input variable of the model. From these data, 19 points were selected for building and training (60%), 7 points for testing (20%), and the remaining ones (7 data points) for validating the developed network.

Feed-forward, back propagated and recurrent layers neural networks were considered for the modeling study. Determination of the number of hidden nodes is one of the main aspects of designing the neural network architecture. Too small a number of hidden nodes limit the ability of the ANN to model the problem, and furthermore such a network may not train well to an acceptable error. On the other hand, too many hidden nodes forced the network for memorizing data rather than learning them for generalization. In this study, the number of hidden nodes was determined by varying the number of nodes, starting with only two hidden nodes, and then adding up to four nodes. For each case, Mean Squared Error (MSE) was computed during training, testing and validating steps.

Table 1 shows the best network design obtained after 20,000 iterations for each case. From this table, it can be found that increasing the number of hidden nodes from 2 to 3 has considerably decreased the MSE% of simulation. Moreover, it can be found that both feed forward and recurrent layers networks can acceptably model the metal purity, ampere efficiency, and also the product rate of the aluminum plant. By increasing the number of hidden nodes from 3 to 4, the MSE of metal purity decreases for both network, but the MSE of ampere efficiency considerably increases. For the product rate, a feed forward network with 4 hidden nodes in the hidden layer has the best accuracy. Conversely, a recurrent layer network with 4 hidden nodes has the least accuracy between all networks.

Table 1 Comparison between the mean square errors

Variable	2 nodes in hidden layer		3 nodes in hidden layer		4 nodes in hidden layer	
	Feed forward	Recurrent	Feed forward	Recurrent	Feed forward	Recurrent
Metal purity (%)	0.0076	0.0078	0.002	0.004	0.0004	0.0014
Ampere Efficiency (%)	0.760	0.897	0.108	0.185	0.205	0.207
Product rate (kg/h)	0.414	0.393	0.407	0.191	0.078	0.908

In Table 2, the number of data points, coefficients, Degree of Freedom (DOF) of the regression and DOF of residual for all cases are presented. From this table, it can be found that for the feed forward and recurrent layers network with 2 nodes in the hidden layer, and also for the feed forward network with 3 nodes in the hidden layer, the number of fitting parameters are less than the number of data points. Therefore, it is obvious that the degree of residual for these models is positive. Hence, the validity of the other networks, i.e. feed forward and recurrent layer networks with 4 nodes in the hidden layer, and also recurrent layer with 3 nodes in the hidden layer, cannot be acceptable.

In Tables 3, 4 and 5, the ANOVA analysis of the presented models for the simulation of output variables i.e., metal purity, ampere efficiency, and production rate are presented. From Tables 3 and 4, it can be found that the feed forward layer network with 3 nodes in the hidden layer has the highest F-value (the mean squared of regression to the mean squared of residuals) for simulating the metal purity and ampere efficiency; therefore, this network can be used for simulating these parameters. Moreover, Table 5 confirms that this network has the lowest error of residuals for the production rate, but due to the higher number of coefficients, its F-value is lower than the networks with 2 nodes in the hidden layer. However, it is still reliable to be utilized for simulating the production rate of the target aluminum plant.The negative value reported for the F-value of other networks is due to the negative value of the DOF, discussed before. As mentioned earlier, these networks cannot be accepted for modeling the Hall-Heroult process.

Parity plots in Figures 3 to 5 show comparisons between the metal purity, ampere efficiency and the production rate versus actual values. As seen from these figures, simulation results obtained from a feed forward-back propagated network with 3 nodes in the hidden layer are in good agreement with the actual data. It should be mentioned that the AAD% of the simulated variables for the metal purity, ampere efficiency, and the production rate were 0.002%, 0.108%, and 0.407%, respectively.

Table 2 Number of coefficients and degree of freedom for different ANN modeling approaches

Variable	2 nodes in hidden layer		3 nodes in hidden layer		4 nodes in hidden layer	
	Feed forward	Recurrent	Feed forward	Recurrent	Feed forward	Recurrent
Number of data	33	33	33	33	33	33
Number of coefficients	19	23	28	37	37	53
DOF* of regression	18	22	37	36	36	52
DOF* of residual	14	10	5	-4	-4	-20

*Degree of freedom

Table 3 Analysis of variance for the metal purity

Variable	2 nodes in hidden layer		3 nodes in hidden layer		4 nodes in hidden layer	
	Feed forward	Recurrent	Feed forward	Recurrent	Feed forward	Recurrent
Error of regression	4.32×10^{-5}	4.09×10^{-5}	3.16×10^{-5}	2.30×10^{-5}	2.42×10^{-5}	1.65×10^{-5}
Error of residuals	9.95×10^{-5}	1.65×10^{-5}	2.36×10^{-6}	-9.1×10^{-6}	-1.37×10^{8}	-2.19×10^{-7}
F-value	4.35	2.44	13.39	-2.53	-175.81	-75.34

Table 4 Analysis of variance for the ampere efficiency

Variable	2 nodes in hidden layer		3 nodes in hidden layer		4 nodes in hidden layer	
	Feed forward	Recurrent	Feed forward	Recurrent	Feed forward	Recurrent
Error of regression	0.596	0.496	0.46	0.344	0.349	0.244
Error of residuals	0.096	0.126	0.0088	-0.011	-0.0006	-0.00649
F-value	6.21	3.93	52.08	-30.34	-581.34	-37.76

Table 5 Analysis of variance for the production rate

Variable	2 nodes in hidden layer		3 nodes in hidden layer		4 nodes in hidden layer	
	Feed forward	Recurrent	Feed forward	Recurrent	Feed forward	Recurrent
Error of regression	254.23	223.52	181.88	135.24	130.023	105.27
Error of residuals	5.24	11.33	18.81	-8.16	-1.53	-18.54
F-value	48.44	19.71	9.67	-16.57	-84.69	-5.68

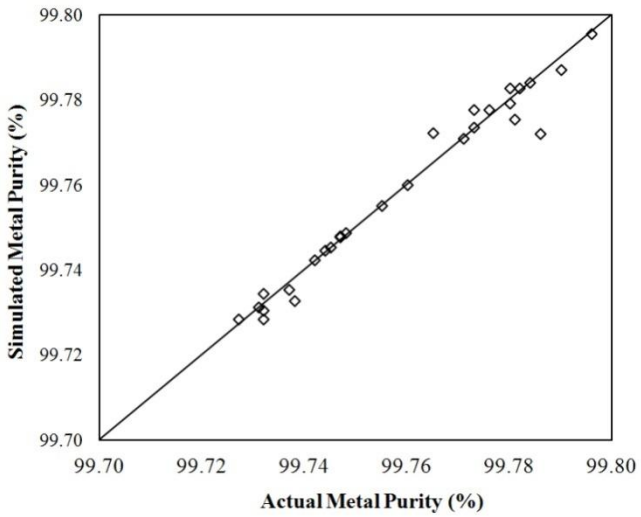

Figure 3 Comparison between the actual data and simulation results for the metal purity obtained from a feed forward model with 3 nodes in the hidden layer

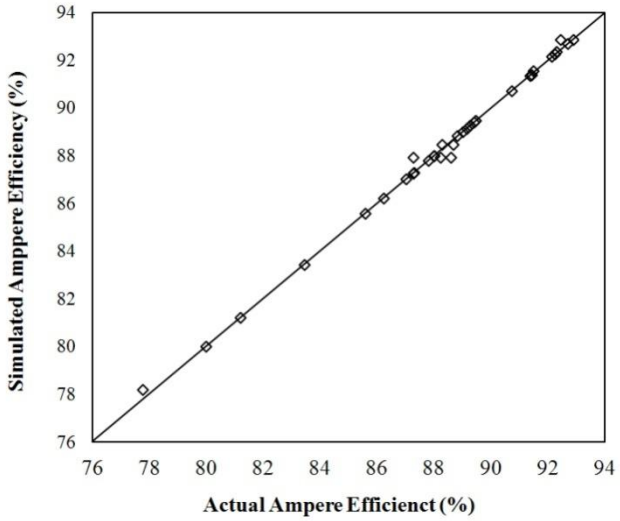

Figure 4 Comparison between the actual data and simulation results for the ampere efficiency obtained from a feed forward model with 3 nodes in the hidden layer

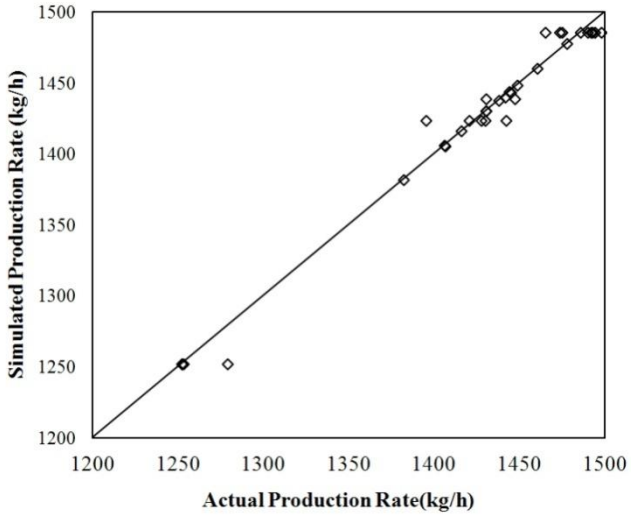

Figure 5 Comparison between the actual data and simulation results for the production rate obtained from a feed forward model with 3 nodes in the hidden layer

After training, testing, and validating the developed ANN model, the developed model was applied to maximize the metal purity and production rate of the target aluminum plant. The optimized decision variables of the Hall-Heroult process i.e., temperature and cell voltage, metal and bath heights, and bath ratio were calculated for all thirty-three data points. For each case, these variables were manipulated to maximize the metal purity and production rate, simultaneously.

Figures 6 and 7 demonstrate the comparison between the actual and optimized metal purity and production rate for all cases. After applying optimal decision variables, the average metal purity and production rate increased by 0.007% and 20kg/h, respectively. Moreover, Figure 7 shows that for the data points with the number of 5, 13 and 29, the optimal production rate is a little lower than the actual one. But, the metal purity of these points shown in Figure 6 confirms that the optimized values are higher than the actual ones; therefore, they can be acceptable.

Additionally, in Figure 8, the ampere efficiency of the aluminum plant for the actual and optimized values is depicted. It is observed that increasing this decision variable (averagely about 0.185%) decreases energy consumption. Therefore, the optimized input variables can also efficiently increase the productivity of the Hall-Heroult process.

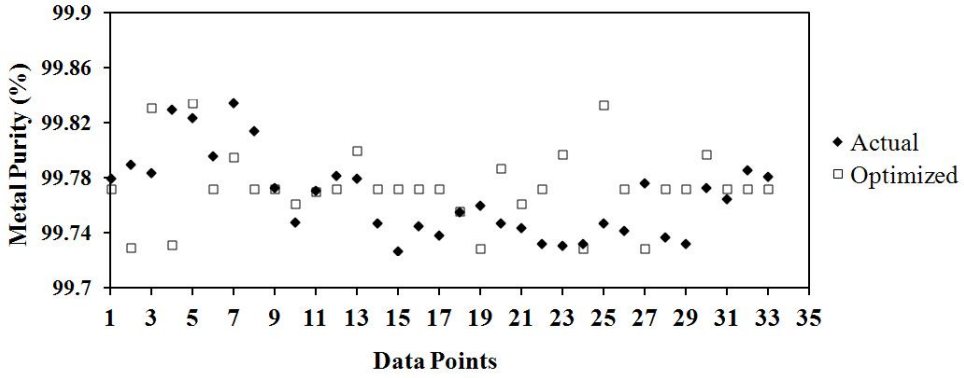

Figure 6 Comparison of the optimized and the actual metal purity

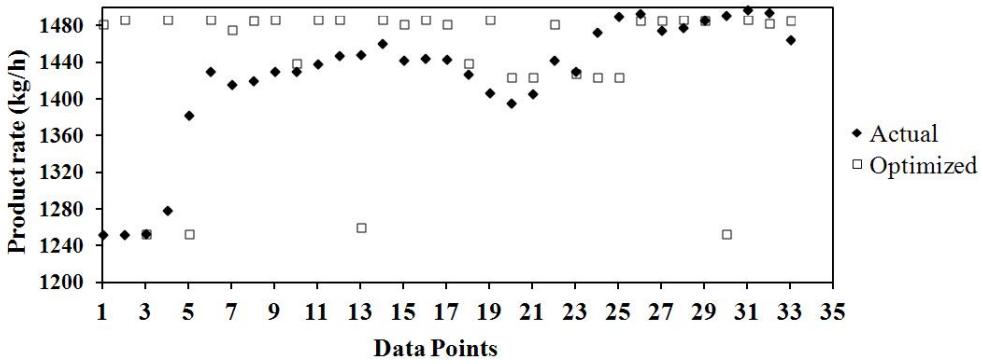

Figure 7 Comparison of the optimized and the actual production rate

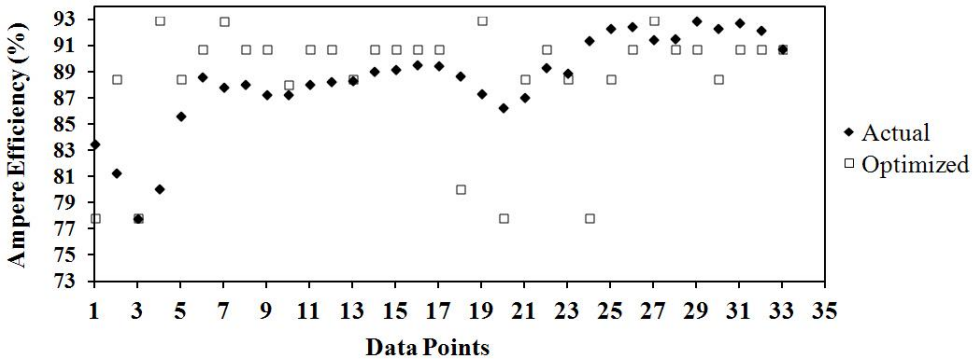

Figure 8 Comparison of the optimized and the actual ampere efficiency

In Figures 9 and 10, the actual temperature and voltage of the cells are compared to the optimized ones. As seen, the optimized cell temperature is lower than the actual value for most of the data points. Therefore, to compensate the decrement of the temperature, the voltage of the cell increases slightly. It can be claimed that decreasing the cell temperature extends the age of the anode in consequence with lowering the carbon consumption. Therefore, decreasing the

cell temperature is another achievement of this research which has an appreciable impact on the economy of the target Hall-Heroult process.

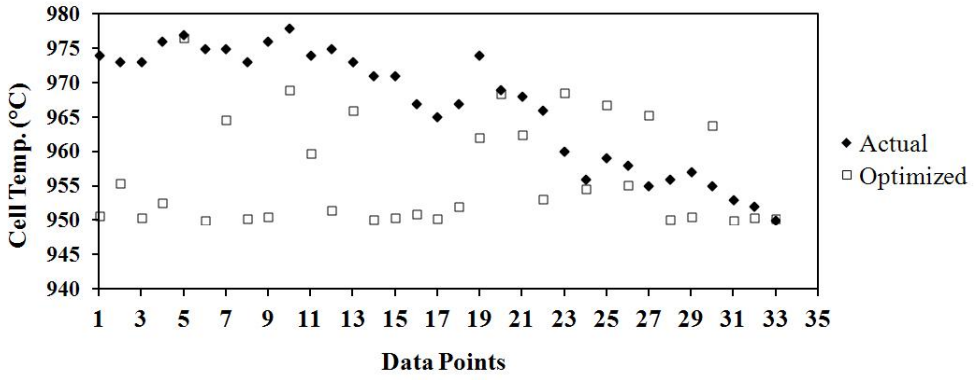

Figure 9 Comparison of the optimized and the actual cell temperature

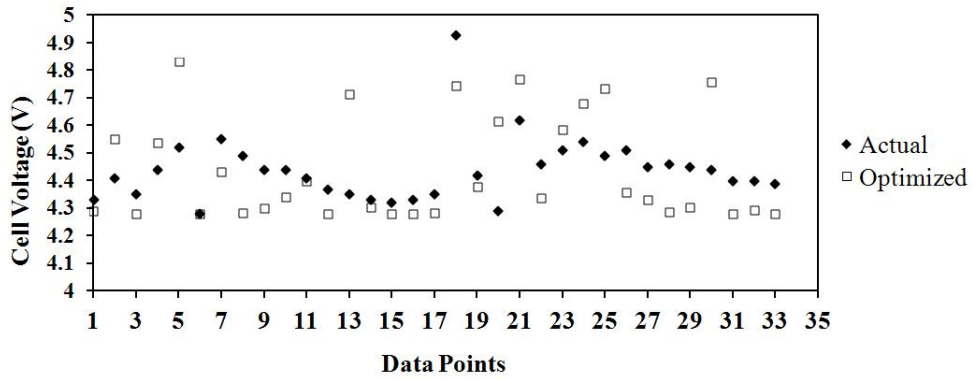

Figure 10 Comparison of the optimized and the actual cell voltage

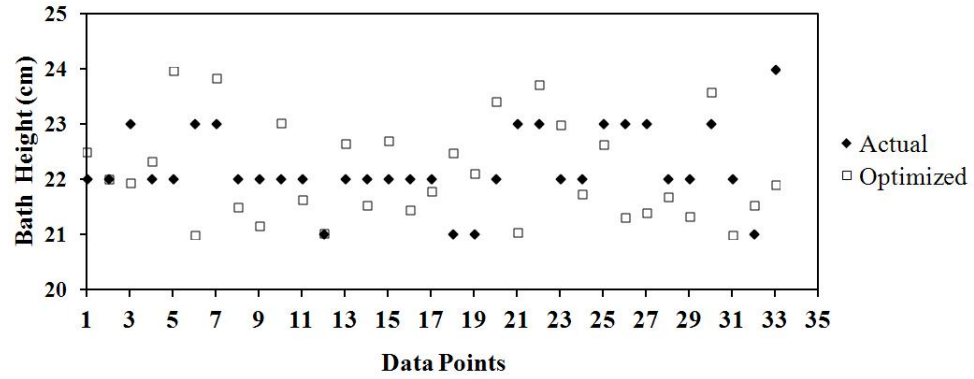

Figure 11 Comparison of the optimized and the actual bath height

In Figures 11, 12 and 13, the optimal values of metal and bath heights and also bath ratio are compared with the actual ones. From the irregular trends of these variables, it can be concluded that to reach to an optimum process conditions, other input variables, such as purity of feed, temperature, and cell voltage should be set meticulously. Therefore, applying an appropriate process model can be beneficial to close the operating condition to an optimum margin.

Figure 12 Comparison of the optimized and the actual metal height

Figure 13 Comparison of the optimized and the actual bath ratio

6. CONCLUSION

Back propagated, feed forward and recurrent layers of Artifical Neural Networks with 2, 3 and 4 neurons in the hidden layer were developed to model the electrolysis cell of an industrial Hall-Heroult aluminum extraction process. The constructed models were trained, tested and validated on the basis of actual data obtained from the plant. These models could simulate the significant output variables of the plant including metal purity, ampere efficiency, and production rate. Results show that the most appropriate ANN model was a feed-forward one, consisting of 3 neurons in the hidden layer, which was capable of simulating the mentioned output variables with the AAD% of 0.002%, 0.108% and 0.407%, respectively. This validated model was then used to carry out a process optimization to maximize metal purity and production rate wherein the decision variables were temperature and cell voltage, metal and bath heights, and bath ratio. Results confirmed that setting these decision variables on the optimal values could approximately increase the metal purity, ampere efficiency and product rate by 0.007%, 0.185%, and 20kg/h, respectively.

7. REFERENCES

Behbahani, R.M., Jazayeri-Rad, H., Hajmirzaee, S., 2009. Fault Detection and Diagnosis in a Sour Gas Absorption Column using Neural Networks. *Chemical Engineering Technology*, Volume 32, pp. 840- 845

Bellos, G.D., Kallinikos, L.E., Gounaris, C.E., Papayannakos, N.G., 2005. Modeling of the Performance of Industrial HDS Reactors using a Hybrid Neural Network Approach. *Chemical Engineering and Processing*, Volume 44, pp. 505- 515

Bhutani, N., Rangaiah, G.P., Ray, A., 2006. First-principles, Data-based and Hybrid Modeling