MODELING OF WET GRINDING OPERATION USING ARTIFICIAL INTELLIGENCE BASED TECHNIQUES

Kishalay Mitra[1] and Mahesh Ghivari

Engineering and Industrial Services, Tata Consultancy Services Limited,
54 B Hadapsar Industrial Estate, Pune 411 013, INDIA

Abstract: The Artificial Intelligence (AI) based modeling techniques applied to the industrial grinding operation of a lead-zinc ore-beneficiation plant to predict the key performance indicators (KPIs) for the circuit. As system identification of the non-linear process is a must in advanced control, AI based techniques are applied to predict the KPIs within some acceptable limits. The nonparametric model for these KPIs is constructed using Feed-Forward Neural Networks (FNN), and wavelet-frames. A well-validated hybrid-model, using physico-empirical methodologies, is used to approximate the actual behaviour of the plant. Merits and demerits of each of these techniques are presented.

Keywords: Process identification, process models, simulation, artificial intelligence, neural network, wavelet.

1.INTRODUCTION

Grinding plays a critical role in most of the ore beneficiation operations in mineral processing plants. As the size of the particles produced while grinding operation becomes the key performance driver for the following separation units, flotation in this case, modeling and thereafter the control of the grinding operation of industrial scale has been a continuous endeavour of the mineral engineers. Over the years, the modeling of grinding operation has attained a reasonable state of robustness (Herbst, *et al.*, 1983; Rajamani, *et al.*, 1991a). Most of these cases, the hybrid path of physical and empirical modeling routes are followed. Research has been carried out on several aspects of single as well as multiple objective optimization and control of industrial grinding operations using hybrid modeling approaches (Rajamani, *et al.*, 1991b). These kind of hybrid grinding models are reported to work really well (Mitra and Gopinath, 2004) if tuned properly with the plant data. But hybrid-modeling approaches have a lot of parameters embedded in empirical correlations. Tuning of these parameters needs a huge data requirement from plant, as this requires measurement in almost every stream. Unfortunately, most of the industrial grinding operations lack in adequate hardware sensors in intermediate streams and have sensors only in the input and final product streams. This makes the tuning process of empirical parameters used in hybrid approach (e.g. grinding, hydrocyclones etc.) extremely difficult. Above this, a multi-variable system identification of the plant operation is very badly needed to control a grinding operation successfully, as control of industrial grinding operation is fairly non-linear in nature and difficult to control. To meet the need of the scenarios stated above, some form of system identification based on data driven modeling procedures come to rescue to deploy control system to operate the plant efficiently. The data based modeling strategies proposed here are based on Artificial Neural Networks (ANN) and wavelets-based networks. Artificial Intelligence (AI) techniques offer

[1]Corresponding author:
Phone: 91-20-4042468; Fax: 91-20-4042399;
E-mail: kishalay.mitra@tcs.com

interesting possibilities for performing the system identification as they provide structures for function approximation with learning capability.

Methods based on neural networks have been proposed as useful tools for process modeling, diagnosis, data rectification and control (Karjala*, et al.*, 1994; Himmelblau*, et al.*, 1993; Pollard*, et al.*, 1992). In identification kind of work, neural networks provide an effective way of initialization. A trained neural network could predict the parameter values called as weights associated with process input and output data and this type of network forms a feed forward identifier. A generic ANN consists of several layers of interconnected neurons. In Feed forward Neural Network (FNN), three types of layers can be distinguished: the input layer (the first layer), the output layer (the final layer) and hidden layers (layers of neuron between the input and output layer). The output of the neuron is determined by functions called as activation functions, which may be non-linear such as sigmoid activation function or squashing function. Each neuron produces a weighted sum of its inputs giving a net result and this net result upon operation by activation function produces the output without any feedback. ANN application development mainly has three phases: the training phase, the testing phase and the users (validation) phase. The training phase determines the weights based on input training pattern, the testing phase calculate output pattern compared against target pattern and the validation phase contains application to an unknown problem. FNN acts as pattern associators and generate a functional relationship that correlates a set of input vectors with its corresponding output vectors.

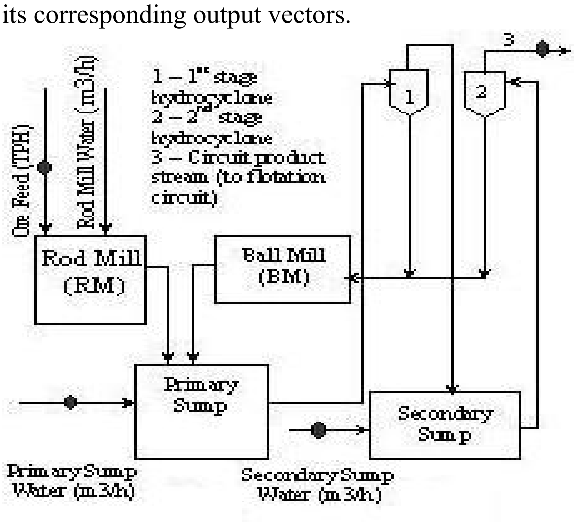


Fig. 1.Industrial grinding circuit

The industrial grinding operation under consideration is a part of two stage (grinding followed by flotation) lead-zinc ore beneficiation process. After crushing in primary and secondary crushers, the ore from the mine is sent to fine ore storage bin from where the fresh ore along with water is fed to the rod mill. The rod mill discharge slurry is mixed with the ball mill discharge slurry in a sump known as the primary sump. Water is added to the primary sump to facilitate the flow of the slurry smoothly within the circuit. The slurry from the primary sump is fed to primary cyclone. The overflow from the primary cyclone goes to another sump, namely secondary sump, where water is added to facilitate the slurry flow further. The mixed slurry from the secondary sump is fed to the secondary cyclone. The underflow product from both cyclones is fed to the ball mill. The overflow from the secondary cyclone is the final product and goes to flotation circuit as feed. The complete circuit configuration is given in Figure 1. In this circuit, only the input and output streams are having the hardware sensors (as shown by black circles in Figure 1) that can indicate the status of key performance indicators (KPI) of the circuit (some properties of slurry at the final product stream) dynamically.

2.FORMULATION

In the grinding circuit presented in Figure 1, three main inputs that are manipulated to control the grinding operation (manipulated variables) are solid stream of raw ore and water streams going to primary and secondary sumps. The circuit has only one output that is secondary cyclone overflow stream. The five KPIs identified for grinding circuit control are throughput (output 1), percentages of three size classes (+150μ, -63μ and -38μ) i.e. output 2, output 3, output 4 respectively and percent solids (output 5) present in the final output stream. Here + sign is used to denote percentage retained whereas – sign is used to denote percentage passing through the given mess size in micron. These are termed as control variables and measured dynamically only at the output stream.

System identification procedure needs data across all operational regimes in which the grinding circuit is operated. For this, the plant has to be run under various possible combinations of input solids and water flow rates. Running the plant over all these possible operating regimes to facilitate data collection is not an affordable task. It disturbs the settings for running the plant in stable mode that incurs a huge loss for this energy intensive process, as the data collection for system identification is a huge time consuming process. For these reasons, input-output data are not collected directly from the plant. Rather they are generated by running numerous simulations from a hybrid (phenomenological and empirical) model of the same industrial operation (Mitra and Gopinath, 2004). This model represents the plant operation very well across all possible operating zones and therefore considered as a very close possible mimic of the plant. Details of this modeling procedure, parameter estimation and plant validation results for this hybrid model can be found in Mitra and Gopinath (2004). Mathematically this hybrid model is a system of differential algebraic equations (DAEs) solved using the DASSL routines. As the data are generated by simulation exercise, it was possible to include another very important KPI for the grinding operation, namely recirculation load that is not generally measured online. This becomes the sixth control variable (output 6) in addition to the earlier five control variables leading to three-input-six-output system identification assignment in hand. Input signals passed to the system in a pseudorandom binary sequence (PRBS) fashion, which ensures coverage of frequencies at wide spectrum. These sets of data gives a relationship between three manipulated variables and six control variables from which the AI based techniques are supposed to churn out the embedded relationship. Two different techniques that are used for system identification here are given below:

(a) FNN: A well-known software package, MATLAB® is used for generating results by FNN technique. 'newff' function of MATLAB® creates a feed-forward back propagation network in which 'trainlm' is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization technique. 'tansig' (non-linear) is taken as input layer transfer function and 'purlin' (linear) is the transfer function for output layer. 'trainlm' function takes the training parameter determined by 'newff' and trains network with input data. Each training iteration is a single representation of all inputs to the network and the network is updated according to the results of presentations.

(b) Wavelet: To characterize the target environment a non-linear nonparametric regression estimator is defined using wavenets as reported in Zhang, *et al.* (1997). The structure under consideration is,

$$Y = f(U) + e \qquad (1)$$

$$U = \{u_1, \cdots, u_n\} \quad \text{and}\ Y = \{y_1, \cdots, y_n\}$$

where Y is the output and $f(U)$ is unknown non-linear function with U is input.

$$\hat{f}(U) = \sum_{i=1}^{N} w_i \psi\ (\alpha_i (U - \beta_i)) \qquad (2)$$

where $\hat{f}$ closely approximates f, w_i is wavelet coefficient, α_i is dilation parameter, β_i is translation parameter and $w_i \in R$, $\alpha_i \in R^d$, $\beta_i \in R^d$, d is the input dimension.

To account for linear regions along with non-linearity linear terms, bias terms are added and above equation is modified to following form,

$$\hat{f}(U) = \sum_{i=1}^{N} w_i \psi\ (\alpha_i (U - \beta_i)) + c^T \Phi + b \quad (11)$$

Mexican Hat is chosen as wavelet basis function.

$$f(U) = \sum_{i=1}^{N} w_i\left((\alpha_i (x - \beta_i))^2 - 1\right) e^{-1/2\left(\alpha_i\left(x - \beta_i\right)\right)^2} + c \cdot x + b \qquad (3)$$

where, $c \in R^d$ is linear coefficients, $b \in R$ is bias terms.

Initialization of the network is done using Akaike's final prediction error criterion (AFPE). Basically, it decides the smallest number of wavelets, which characterizes the target environment.

$$E_{AFPE}\left(\hat{f}\right) = \frac{1 + m_q/m}{1 - m_q/m} \frac{1}{2m} \sum_{k=1}^{n} \left(\hat{f}\left(x_k\right) - y_k \right)^2 \quad (4)$$

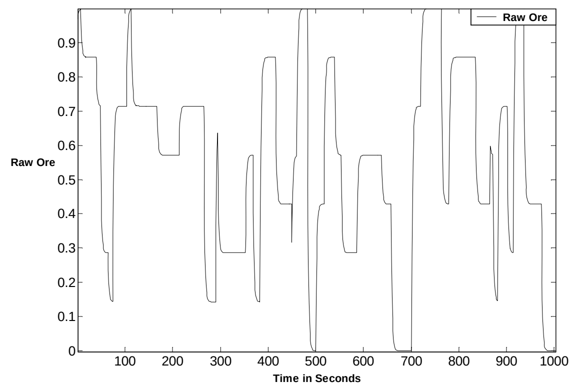
$$m_q = M\ (q + 2) + q + 1 \qquad (5)$$

m_q is Number of parameters in the estimator, $(x_k, y_k) \in O_1^Z$ is training data of size Z, M is number wavelets in network.

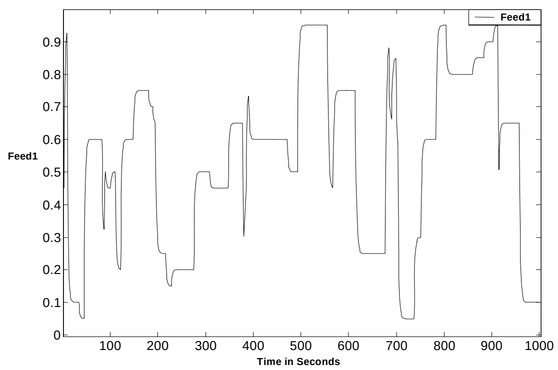
Backward elimination algorithm is used to choose best wavelet regressors as reported in Zhang, *et al.* (1997).

3.RESULTS AND DISCUSSIONS

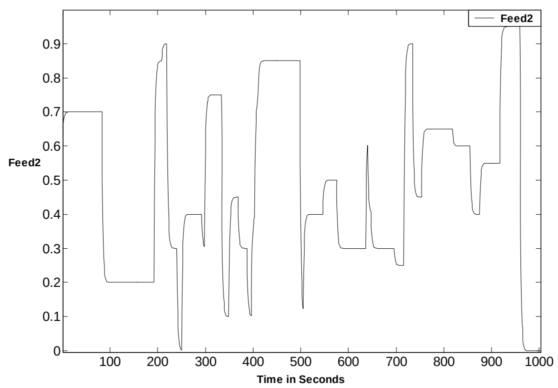
The data for KPI values for grinding circuit final output stream are generated using hybrid-grinding dynamic model of Mitra and Gopinath (2004) by applying different input patterns to it.



(2a)



(2b)



(2c)

Fig. 2. Excitation patterns for (a) raw ore to rod mill (b) water to primary sump (Feed1) and (c) water to secondary sump (Feed 2)

The input patterns of the MVs were randomly generated so as to cover the all-possible frequencies that may be involved in actual plant operation, inline with so-called PRBS (as shown in Figure 2). Excitation is given to three MVs within their respective operating bounds except rod mill water, which is kept constant at a pre-specified value. Part of the data set generated from model is used in training phase and part of it is used for testing and validation phase combined together. This is essential in a sense to create a sort of noise in the system and thus test the model for an unseen data.

3.1 Feed-forward Neural Network

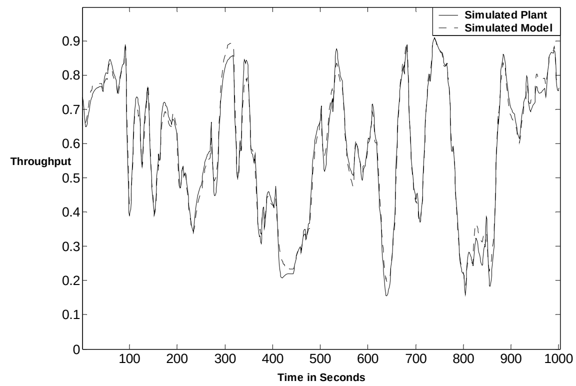
In case of FNN, the tuning parameters are number of hidden nodes, iterations, training data size and unfolding of the input parameters. Table 1 shows comparative study of different tuning parameters in testing phase. Data in row 1 in table 1 (T1R1) and T1R2 show the results based on number of training data size. Here the mean square error (MSE) seems to have increased a little since less variability in data. Data size of 10K seems to contain enough information. T1R1 to T1R5 show the effect of hidden nodes where 3 numbers of nodes found to yield good results. T1R6 to T1R9 tell us how many numbers of input data unfolding is required to attain the minimum MSE. It can be seen that MSEs are reduced drastically for unfolding of 20 depicting the dynamics present in the system. In T1R10, T1R11 learning rate has been changed to assess the performance. Lower learning rate leads to more error as well as computation time. The computation time also increases due to lower converging rate. It is found that the system is well captured at a learning rate of 0.009, unfolding of 20 and 3 hidden nodes. As can be seen from Figure 3a to 3d FNN has only partially captured sharp transients.

Table 1a: Parameters for FNN

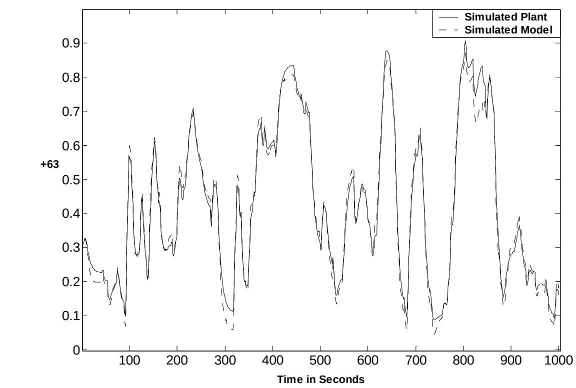
Sr. No.	Cycles	Hidden Nodes/ layer	Learning Rate	Training Data	Testing Data	Unfoldings Of inputs
1	20	3/1	0.009	10K	10K	0
2	20	3/1	0.009	30K	10K	0
3	20	5/1	0.009	10K	10K	0
4	20	10/1	0.009	10K	10K	0
5	20	2/1	0.009	10K	10K	0
6	20	3/1	0.009	10K	10K	5
7	20	3/1	0.009	10K	10K	10
8	20	3/1	0.009	10K	10K	20
9	20	3/1	0.009	10K	10K	25
10	20	3/1	0.008	10K	10K	20
11	20	3/1	0.01	10K	10K	20

Table 1b: Results of simulations based on FNN

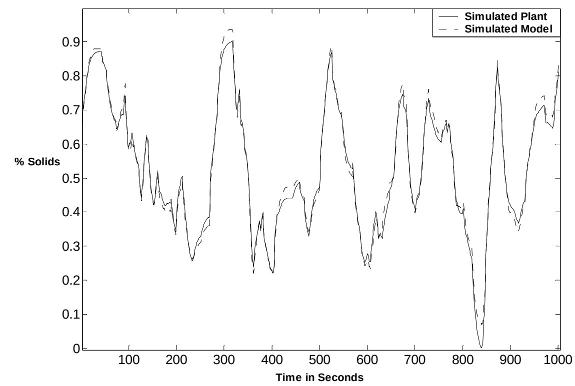
Sr. No.	MSE1	MSE2	MSE3	MSE4	MSE5	MSE6	Time in Seconds
1	0.8816	0.0024	0.4063	0.1859	0.1361	3.9269	46
2	0.9668	0.002	0.4143	0.1938	0.1488	5.1148	133
3	0.8831	0.002	0.4026	0.1842	0.1365	3.6931	71
4	0.8815	0.002	0.4087	0.1881	0.1391	3.6946	152
5	0.9705	0.0024	0.4178	0.1949	0.1538	5.3928	35
6	0.3264	9.57E-04	0.1241	0.0567	0.046	0.6252	113
7	0.1429	6.45E-04	0.0496	0.0229	0.0148	0.2137	199
8	0.0788	5.16E-04	0.0252	0.012	0.0092	0.0688	393
9	0.0716	5.30E-04	0.0226	0.0109	0.0109	0.1421	559
10	0.0713	8.53E-04	0.022	0.0108	0.0091	0.4038	396
11	0.0727	5.48E-04	0.0221	0.0107	0.0111	0.106	392



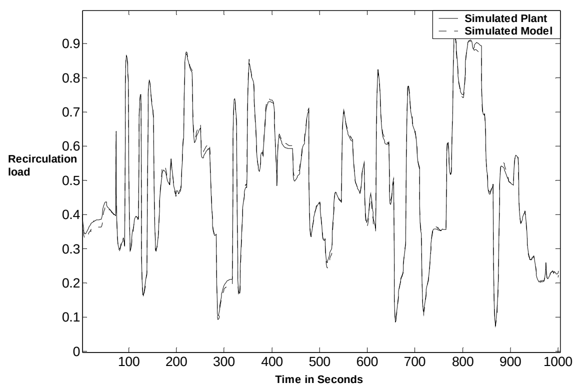
(3a)-Throughput



(3b)- size fraction on + 63µ



(3c)- % solids



(3d)- Recirculation load

Fig. 3. Output responses for FNN.

3.2 Wavelet

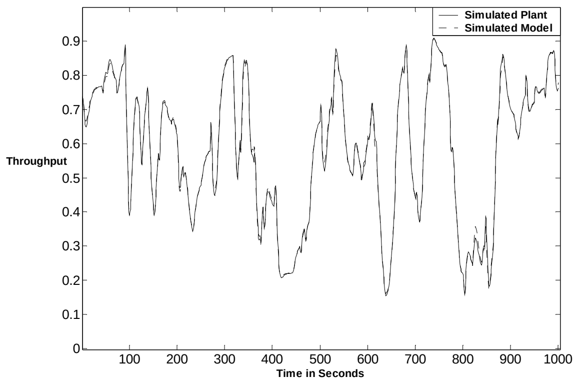
Comparisons with respect to tuning parameters are given in Table 2. From T3R1, T3R2, T3R7, T3R8 it is seen that a single wavelet node is required to do the required modeling as with one wavelet mode, the MSE presented is the least. T3R2 to T3R4 show that numbers of observations for repeating patterns are not large, resulting in less number of wavelet nodes. It is evident from T3R4 to T3R6 that not many numbers of levels are required to cover all regions of interest and a single level is sufficient. T3R7, T3R9, T3R10, T3R11 readings give an idea about the dynamics present in system. More number of unfolding leads to increased computation time and little change in MSE. For output number 2, from T3R14, T3R15, number of iterations has considerably reduced the MSE at the cost of computation time while in case of output number 3 and 4, from T3R16 to T3R19, MSE has shown less

Table 2a: Tuning parameters for wavelet

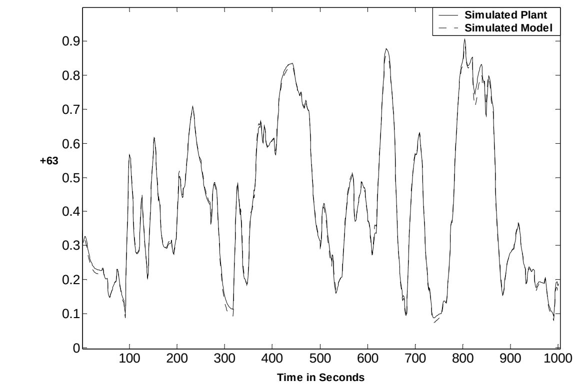
Sr. No.	Cycles	Output No.	Min. Obs.	Level	Wavelet Nodes	Training Data	Testing Data	Unfolding Of inputs
1	1/20	1	100	3	10	10K	10K	20
2	1/20	1	100	3	1	10K	10K	20
3	1/20	1	500	3	1	10K	10K	20
4	1/20	1	1000	3	1	10K	10K	20
5	1/20	1	1000	2	1	10K	10K	20
6	1/20	1	1000	1	1	10K	10K	20
7	1/20	1	1000	1	1	10K	10K	10
8	1/20	1	1000	1	3	10K	10K	10
9	1/20	1	1000	1	1	10K	10K	15
10	1/20	1	1000	1	1	10K	10K	25
11	1/20	1	1000	1	1	10K	10K	30
12	1/20	1	1000	1	1	10K	10K	35
13	1/20	1	1000	1	1	30K	10K	30
14	1/20	2	1000	1	1	10K	10K	30
15	20/20	2	1000	1	1	10K	10K	30
16	1/20	3	1000	1	1	10K	10K	30
17	20/20	3	1000	1	1	10K	10K	30
18	1/20	4	1000	1	1	10K	10K	30
19	20/20	4	1000	1	1	10K	10K	30
20	1/20	5	1000	1	1	10K	10K	30
21	1/20	6	1000	1	1	10K	10K	30

Table 2b: Results for wavelet

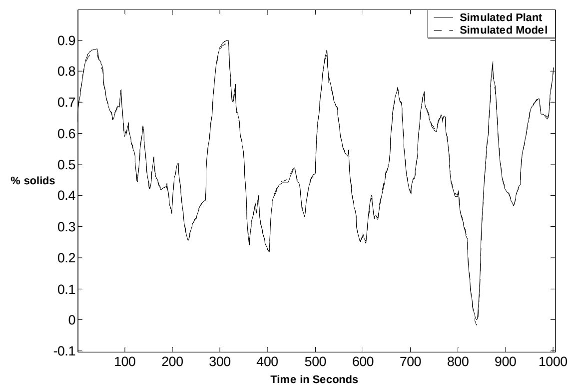
Sr. No.	Output No.	MSE	Time taken for computation in Seconds
1	1	0.013017	1400
2	1	0.010795	133
3	1	0.010795	664
4	1	0.010795	649
5	1	0.010795	154
6	1	0.010795	29
7	1	0.076924	16
8	1	0.076924	168
9	1	0.022936	21
10	1	0.008395	35
11	1	0.007958	45
12	1	0.007865	59
13	1	0.007826	108
14	2	0.000837	35
15	2	0.000565	439
16	3	0.004403	33
17	3	0.003827	488
18	4	0.002638	35
19	4	0.002405	492
20	5	0.001021	83
21	6	0.150806	31



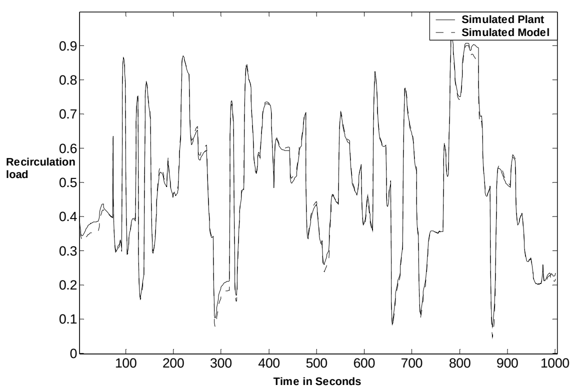
(4a) -Throughput



(4b) - size fraction on + 63μ



(4c) - % solids



(4d) - Recirculation load

Fig. 4. Output responses for wavelet.

change over the number of iterations, even at the cost of computation time. Output number 1,5 and 6 are converged in single iterations since increase in cross generalization error has terminated the number of iterations. T3R11 and T3R13 show the results based on number of training data size required. It can be inferred that additional number of data set does not provide any useful information. It is observed that the different tuning parameter cases applied for output 1 are approximately applicable for the rest of cases. Hence only different cases pertaining to output 1 are evaluated. From The simulation results in Figure 4a to 4d show that the localized effects (sharp transients) are also captured. The best tuning parameters obtained are unfolding number of inputs 30, level 1, number of minimum observations 1000.

All results reported here are normalized to maintain the data secrecy agreement. For showing the FNN and wavelet model predictions, trends for only 4 CVs are shown (Figure 3 and 4) for the shake of brevity. The comparative study of the two different AI based techniques while modeling the industrial grinding operation can be summarized as:

(1) Wavelets are able to capture local behavior (transients) as well as global (mean) behavior while neural networks have shortcoming of not capturing transients at local region

(2) Neural network algorithm is run for Multiple input multiple output (MIMO) system while wavelets are for Multiple input single output (MISO). This is due the basic difference in the methodologies they use for function approximation. ANNs use the object function constructed as a sum of squares of the errors between the measured values and estimated values of the output and is minimized to get the weights of the network. The wavelet networks adjust the dilation and translation parameters in addition to the coefficients for the nodes, which poses a challenge for handling multiple outputs in a single topology.

(3) Both techniques are able to represent similar dynamics based on number of unfolding required.

(4) Wavelet offers more accuracy in terms of MSE than ANN for the problem under consideration however is static in nature and requires dynamics to be added in the form of regressions of the inputs.

These kinds of AI technique based models can be used as softsensors for grinding operation. Once a softsensor of this nature is present, one can make use of the same for grinding operations optimization and control and thereby running the energy expensive grinding operation in optimized fashion. This cannot only enhance the operation but also imparts stability into the system leading to various practical industrial tangible and intangible benefits reported by Mitra and Gopinath (2004).

4.CONCLUSIONS

Artificial Intelligence techniques are applied for modeling of an industrial Pb – Zn grinding operation. All the six KPIs were modeled using three inputs. Both techniques found to be good at fitting the data. This gives a confidence in applying data based modeling techniques to industrial grinding operations. Additionally it saves time on building first principles models, find the required physical parameters and validate the models. Whenever a control strategy is implemented for a particular process, largely its success depends upon the online measurements, reliability of the data and continuous availability of the data. Most controllers fail because of the former mentioned reason irrespective of complex and powerful control strategy devised for the class of the problem. Thus, AI based models tend to offer an alternative to online measurements acquired from hardware sensors and can be used as soft sensors. AI based softsensors, when used in addition with optimization and control methodologies, are capable of reaping tremendous tangible and intangible benefits to industrial grinding operation.

5.ACKNOWLEDGEMENT

Authors acknowledge Mr. Shivram Kamat and Mrs. Jessy Smith for their inputs on ANN and wavelet and support by Dr. Gautam Sardar in course of this work.

6.REFERENCES

Carrier, F. J. and G. Stephanopoulos (1998). Wavelet-Based Modulation in Control-Relevant Process Identification, AIChE J, **44(1)**, 341-359.

Daubechies, I., (1992). Ten Lectures on Wavelets, CBNS-MSF regional series in applied mathematics, Philadelphia: SIAM.

Herbst, J. A., M. Siddique, K. Rajamani and E. Sanchez (1983). Population Based Approach to Ball Mill Scale-up: Bench and Pilot Scale Investigations, Trans AIME, **272**, 1945-1954.

Himmelblau, D. M. and J. C. MacMurray (1993). Modeling and Control of a Packed Distillation Columns using Artificial Neural Networks, Comp. Chem. Eng.,

Karjala, T. W. and D. M. Himmelblau (1994). Dynamic Data Rectification by Recurrent Neural Networks vs. Traditional Methods, AIChE J, **40(11)**, 1865-1875.

Mitra, K. and R. Gopinath (2004). Multiobjective Optimization of an Industrial Grinding Operation Using Elitist Nondominated Sorting Genetic Algorithm, Chem. Engg. Sci., **59**, 385–396.

Pollard, J. F., M. R. Broussard, D. B. Garrison and K. Y. San. (1992). Process Identification using Neural Networks, Comp. Chem. Eng., **16(4)**, 253-270.

Rajamani, R. K. and J. A. Herbst (1991a). Optimal Control of a Ball Mill Grinding Circuit-I. Grinding Circuit Modeling and Dynamic Simulation, Chem. Eng. Sci., **46(3)**, 861-870.

Rajamani, R. K. and J. A. Herbst (1991b). Optimal Control of a Ball Mill Grinding Circuit-II. Feedback and Optimal Control, Chem. Eng. Sci., **46(3)**, 871-879.

Zhang, Q., (1997). Using Wavelet Network in Nonparametric Estimation, IEEE Trans. on Neural Networks, **8(2)**, 227 – 236.