73

Solids Residence Time Distribution in a Three-Phase Bubble Column Reactor: An Artificial Neural Network Analysis

V.K. Pareek[1], R. Sharma[2], C.G. Cooper[3] and A.A. Adesina*,[3]

[1] *Department of Chemical Engineering, Curtin University of Technology, Perth, WA 6085, Australia*

[2] *EPIN System Pvt. Ltd., 14, Uniara Garden, Moti Doongri Road, Jaipur, 302 004, India*

[3] *Reactor Engineering & Technology Group, School of Chemical Sciences & Engineering, University of New South Wales, Sydney, NSW, 2052, Australia*

Abstract: Residence time distribution (RTD) study of solids in a three-phase pilot-scale bubble column photoreactor has been carried out in order to provide data for the development of an artificial neural network model usable for process optimisation. The experimental data indicated that the RTD of solids was a complex nonlinear function of gas and liquid velocities as well as the contacting pattern (co-current and countercurrent flow of gas and liquid). In this study, the solid particle RTD data were modeled using feed forward artificial neural networks (ANN). The networks were trained with 250-sets of input-output patterns using back-propagation algorithm. The trained networks were tested using 50-sets of RTD data previously unknown to the networks. Out of several configurations, a 3-layered network with 6-neurons in its hidden layer yielded optimal results with respect to the validation data. The optimal model and empirical data exhibited good agreement with a correlation coefficient of 0.995.

INTRODUCTION

Due to their excellent heat and mass transfer characteristics, three-phase bubble column reactors (BCRs) are the preferred reaction vessels for many chemical process operations including production of fine chemicals, coal liquefaction, wastewater treatment, photocatalytic and fermentation processes. The mixing of solids in a three-phase reactor is critical to its overall performance [1-3] and this is especially so in a photocatalytic bubble column reactor where the catalyst particles exposure to light and its distribution within the vessel impacts upon the local volumetric rate of photon absorption (LVRPA). Until recently, lumping approaches based on axial dispersion model (ADM) have been used to describe the residence time distribution (RTD) of solids within three-phase reactors. Being a single-parameter model, ADM is relatively easy to use but does not account for other important pathological phenomena (channeling, by-passing, dead-zones, etc) in the BCR. In reality, solid particle dynamic behaviour in a BCR is a multiple-input-multiple-output (MIMO) system and in the absence of detailed mechanistic information, it is difficult to secure reliable and adequate model useful for predictive, control and optimization purposes.

Kiared *et al.* [2] proposed a Cross Flow Multistage Stirred Reactors (CFMSR) model to describe solid phase RTD pattern in three-phase reactors. Their model was based on experiments conducted in a 10 cm ID bubble column

reactor. Using Radioactive Particle Tracking (RPT), they identified two different regions in the bubble column reactor, namely; an ascending wake region and a descending emulsion phase. In the CFMSR model, each region was approximated by a tank-in-series model with allowance for inter-phase mass (solid) exchange (constant mass transfer coefficient) between both regions. Pareek *et al.* [3] performed RTD experiments on Aldrich titania particles in an 18-L pilot scale reactor. They studied both co-current and countercurrent contacting of gas and liquid. Interestingly, the RTD curves from the co-current operation showed bimodal behaviour at relatively higher gas velocities. The solids mixing process of this operation was modeled as a parallel arrangement of tanks-in-tandem and a plug-flow reactor with inter-branch solid exchange. On the other hand, the RTD pattern for counter-current operation was modeled as a series of stirred-tank reactors with recycle stream. The analysis revealed an evolution of different solid flow patterns even at constant fluid flow rate as evidenced from the time-dependent behaviour of the intensity function.

Prediction of the solid phase dispersion is a challenging task since it depends on a large number of parameters such as gas and liquid velocities, particle size distribution, physical and transport properties of gas and liquid as well as reactor size. Although the empirical models developed previously gave reasonably good predictions with mean square error between 3-10 %, the lack of a comprehensive framework in which all these factors are explicitly related to reactor performance indices limits their application for optimization and control purposes.

In recent years, the concept of artificial neural nets (ANN) has received wide popularity in many areas of

*Address correspondence to this author at the Reactor Engineering & Technology Group, School of Chemical Sciences & Engineering, University of New South Wales, Sydney, NSW, 2052, Australia; Tel: +61-2-9385 5268; Fax: +61-2-9385 5966; E-mail: a.adesina@unsw.edu.au

chemical engineering [4]. The ability of ANN to recognize and reproduce cause-effect relationships through training, for complex multiple input-output systems, makes them efficient to represent complex systems [5]. Roedemerck *et al.* [6] have employed ANN to determine optimal catalyst composition for the oxidative dehydrogenation of propane. Gunay & Yildrim [7] also developed and optimized alumina-supported Pt-Co-Ce catalyst for selective CO hydrogenation in an H$_2$-rich stream using ANN modeling while Molga [8] has presented a generalised neural network approach for the modeling of catalytic reactors. In what follows, we demonstrate the deployment of the ANN technique to analyse and comprehensively model RTD data obtained from an 18-litre pilot scale reactor. A set of 300-RTD data points was used to train and test the neural nets [3]. The input layer of the neural nets consisted of liquid velocity, gas velocity, time and contacting scheme, and the output layer contained one neuron – the RTD function, *E*(*t*).

EXPERIMENTAL DETAILS

Apparatus

The system contains two concentric cylinders – an outer chamber ID=20 cm, and the inner cylindrical lamp, OD=5 cm [3]. The liquid and gas velocities employed ranged from 0.03-0.13 cm s^{-1} and 0.05-0.5 cm s^{-1} respectively. This ensured that all operations were carried out in the pseudo-homogeneous flow regime and minimised the requirement for longer disengaging height (present height is about 25% of the liquid column, 70 cm, in the vessel) at higher liquid and gas flow rates as well as preventing damage to the fragile co-axial UV quartz lamp in the reactor. Iwaki magnetic pump was used for liquid delivery while electronic mass flow controllers metered all gas flow rates. Air from the flow controller was passed upwards through the moving column of water *via* a 70μm stainless mesh distributor. For each pair of gas/liquid velocity combination, a 1g pulse of commercial titania powder was instantaneously injected at the column base and the concentration and size distribution of the particles at the outlet were monitored using Malvern Mastersizer/E. Both co-current and counter-current combinations of gas-liquid flow modes were studied.

Materials and Methods

Titania (>99% anatase) was obtained from Aldrich Chemicals and used as supplied in all runs. The nominal particle size of titania powder in aqueous slurry was between 0.2 to 200 microns. Ordinary domestic water was used as liquid phase. Air at room temperature was used as the feed gas for the reactor.

ARTIFICIAL NEURAL NETS (ANN)

The human brain is a natural example of the performance of a MIMO system and artificial neural net is an engineering attempt to mimic the way this versatile organ works. In human brain, billions of neurons are interconnected to process a variety of complex information. However, a computational neural network is much simpler than the human brain and consists of only a few dozens of neurons. A neural net consists of simple processing units called neurons. A neuron is an entity capable of receiving and sending signals, and it is simulated by means of software algorithms on a computer. A typical artificial neural net is shown in Fig. (**1**). In general, a

neural net is parallel interconnected structure consisting of, (1) input layer of neuron (independent variables), (2) a number of hidden layers, (3) and output layer (dependent variables). The number of input and output neurons is determined by the nature of the problem. The hidden layers may be considered as a processing or features-detecting device.
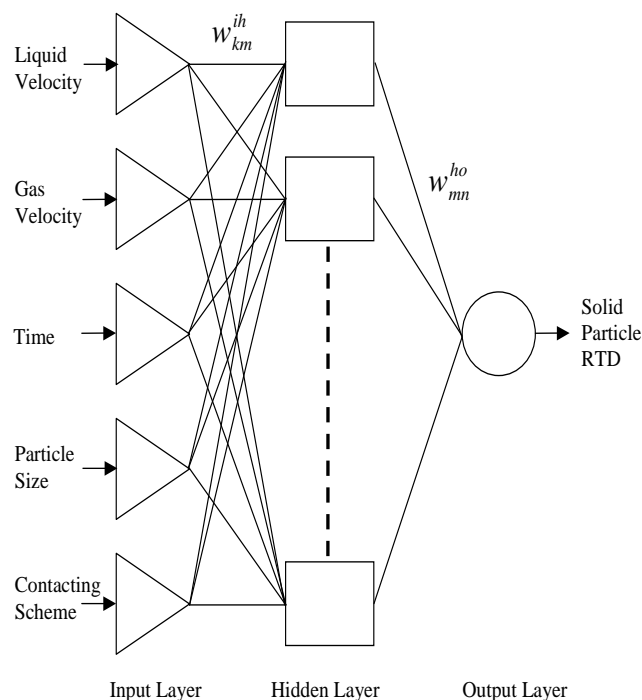

Fig. (1). Schematic diagram of the computational neural network.

The input layer acts like a feeder and it directly transmits the information to the hidden layer. The inputs to hidden and output-layers are calculated by performing a weighted summation of all the inputs received from the preceding layer. Generally, the output from the hidden layer is calculated by using a transfer function. The function should be monotonic, continuously differentiable and bounded. Among others, the logistic sigmoid transfer function is the most widely used [4, 7]:

$$f(x) = \frac{1}{1 + e^{-x}} \qquad (1)$$

Although it is possible to have several hidden layers in an ANN, according to the universal approximation theory, a network with a single hidden layer with a sufficiently large number of neurons can interpret any input-output structure [9, 10]. As a result, we have used only one layer of hidden neurons. The actual number of neurons in the hidden layer is determined by the desired accuracy in the neural predictions. Hence, it may be considered as a parameter for the neural net design. Mathematical expressions used for each of the layers in the feed-forward ANN model are summarised in Table **1**.

MODEL DEVELOPMENT

The development of an ANN model generally consists of three steps: (i) generation of input-output data for training and validation purposes; (ii) training of the neural net to optimise model parameters and (iii) the testing of neural nets. During the training step the net is exposed to a certain num-

Table 1.   Summary of Input-Output of Neurons in a 3-Layered Feed Forward Neural Network

Layer	Neuron identity	Input to the neuron	Output from the neuron
Input	i	x_i	$y_i = x_i$
Hidden	h	$x_h = \sum_{k=1}^{N_i} w_{ih}\, y_i$	$y_h = \dfrac{1}{1 + \exp(-x_h)}$
Output	o	$x_o = \sum_{m=1}^{N_h} w_{ho}\, y_h$	$y_o = x_o$

ber of input-output patterns and an objective function is used to minimise the errors in predicted and target values. Finally, in the testing step, the net is exposed to the patterns unfamiliar to it and the accuracy in the predictions is evaluated.

Data Collection

Since an artificial neural net learns by examples, it is essential to have a large number of data sets to enable a valid net training. A representative training data set should include the cause-effect relationship between the input and output variables. In this case, the sets of training and testing data were taken from RTD experiments performed in the 18-L pilot scale reactor. Typical data patterns are shown in Fig. (**2**). Out of the 300 data points selected for this study, 250 points were chosen for training the neural nets and the remaining 50 data points were used in the validation step.

Training of Artificial Neural Nets

The training of artificial neural nets (ANN) is essentially an optimisation process, where an error function is minimised

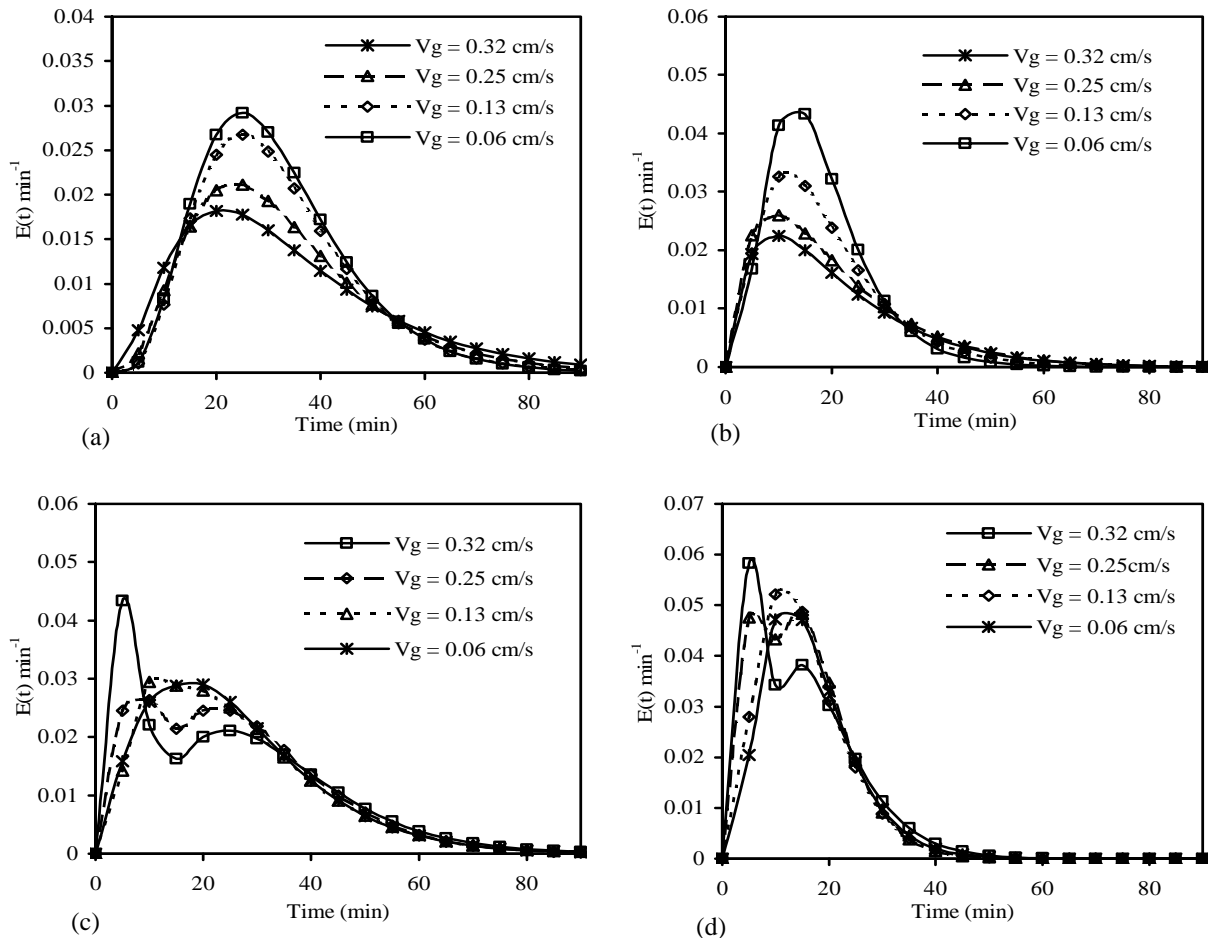RTD OF SOLIDS IN THREE PHASE REACTORS

Fig. (2). Typical residence time distribution curves. (a) Countercurrent flow of gas and liquid ($V_L = 0.06$ cm s^{-1}), (b) Countercurrent flow of gas and liquid ($V_L = 0.13$ cm s^{-1}), (c) Co-current flow of gas and liquid ($V_L = 0.06$ cm s^{-1}), and (d) Co-current flow of gas and liquid ($V_L = 0.13$ cm s^{-1}).

by changing the neural net weights. The most widely used error function is the total sum of squared-error defined as [11]:

$$e = \frac{1}{2} \sum_{j=1}^{N_s} \sum_{n=1}^{N_o} (Y_{o,j} - T_{o,j})^2 \qquad (2)$$

where, the indices j and o refer to a pattern and output neurons respectively.

Following exposure of the net to a new training pattern, it calculates the output using the input variables and the error given by equation (2). If the computed error is less than a pre-specified value, the training stops, otherwise the weights of neural net are adjusted and all the calculations are performed again. In the present case, the back-propagation algorithm was used to update the weights [12]. In this algorithm, for every input-output pattern, a forward-pass is used to calculate the output and a backward-pass is used to adjust the weight. Firstly, the weights connecting hidden and output layers are adjusted using generalised delta rule [11]:

$$w_{ho}(p + 1) = w_{ho}(p) + \Delta w_{ho}(p) \qquad (3)$$

$$\Delta w_{ho}(p) = \beta \sum_{j=1}^{N_s} \left(y_{o,j}(1 - y_{o,j})(y_{o,j} - T_{o,j}) y_h \right) \\ + \alpha \Delta w_{ho}(p-1) \qquad (4)$$

α and β are momentum and learning rate parameters (generally, their values are set between 0.6 an 0.9), index p refers to pattern number or iteration number. Similarly, the weights for the connections between input and hidden layer are updated:

$$w_{ih}(p + 1) = w_{ih}(p) + \Delta w_{ih}(p) \qquad (5)$$

$$\Delta w_{ih}(p) = \beta \sum_{j=1}^{N_s} \left(y_{h,j}(1 - y_{h,j}) \sum_{n=1}^{N_o} \delta_{o,j} w_{ho}(p) \right) \\ + \alpha \Delta w_{ih}(p-1) \qquad (6)$$

where, $\delta_{o,j} = y_{o,j}(1 - y_{o,j})(y_{o,j} - T_{o,j})$ is error gradient for the output layer.

In this study, over 40 neural net models were trained and tested using the 300 input-output patterns. As shown in Fig.

(**1**), all the models developed consisted of three layers. The first layer was input layer and in this case it had 5–neurons (liquid velocity, gas velocity, particle size, time, and contacting scheme: +1 if co-current, -1 if countercurrent). The second layer was hidden layer and it contained variable number of neurons (2-20). The third layer was the output layer with one neuron (the RTD of solids).

Selection of Optimal Net Configuration

Several net configurations with one hidden layer were trained using the input-output patterns. The trained nets were used to predict the RTD for 50 input data previously not known to the nets. Following error criteria were used to assess the model predictions:

$$C_p = \frac{\sum_{j=1}^{N_s} (E_{p,j} - \overline{E}_{p,j})(E_{a,j} - \overline{E}_{a,j})}{\sqrt{\sum_{j=1}^{N_s} (E_{p,j} - \overline{E}_{p,j})^2 \sum_{j=1}^{N_s} (E_{a,j} - \overline{E}_{a,j})^2}} \qquad (7)$$

$$RMSE = \sqrt{\left(\sum_{j=1}^{N_s} (E_{p,j} - E_{a,j})^2 \right) \Big/ N_s} \qquad (8)$$

$$MAPE = \frac{1}{N_s} \sum_{j=1}^{N_s} \left(\left| E_{p,j} - E_{a,j} \right| \Big/ E_{a,j} \right) \times 100 \qquad (9)$$

where, the index 'j' refers to validation pattern number and $\overline{E}$ to average value of RTD data points.

A summary of various error measures, used to compare the net efficiency, is shown in Table **2**. Mean absolute percentage error (MAPE) in all the cases was less than 10 % while correlation coefficients were generally greater than 0.90, indicating that all the configurations studied in this work could efficiently represent the experimental data. However, a neural net with 6-hidden neurons was found to yield the most optimal results. The predicted RTD with this configuration has been plotted against the actual RTD in Fig. (**3**). The correlation coefficient between the two data sets was 0.995, which is indicative of an excellent agreement between the predicted and actual data. The value of MAPE for this network configuration was 3.43, indicating the optimality of this configuration.

Table 2. Prediction Errors and Correlation Coefficients

N^h	C_p	RMSE	MAPE (%)
2	0.893	0.07	8.73
4	0.986	0.05	5.66
6	0.995	0.03	3.43
8	0.984	0.04	4.23
10	0.991	0.03	3.56
12	0.990	0.03	3.83
Phenomenological models [3]	0.912	0.06	6.41

RMSE: root mean square error.
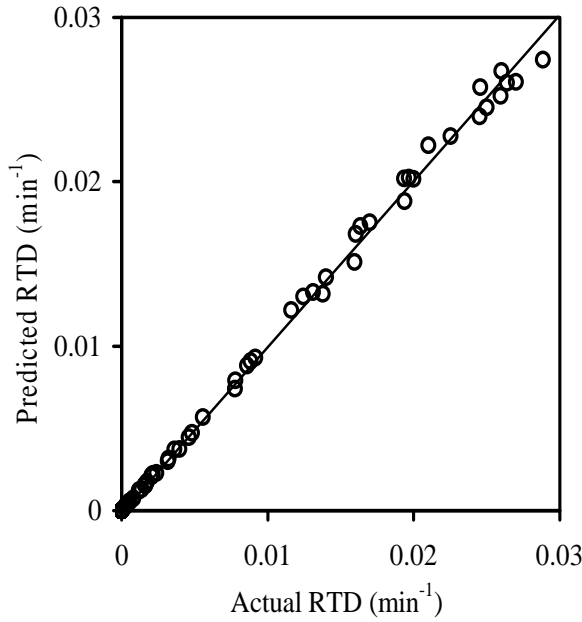MAPE: mean average percentage error.

Fig. (3). Comparison between ANN predictions and actual RTD data.

Comparison with Phenomenological Models

In our previous paper, we had reported two phenomenological models for the co-current and countercurrent flow of gas and liquid [3]. As shown in Figs. (**2a**) and (**2b**), the RTD curves for countercurrent operation showed the usual unimodal behaviour. Therefore, the countercurrent RTD of solid particles was modeled as a series of stirred tanks with a recycle stream. The relevant expression for the E-curve is given by [3]:

$$E(t) = \frac{1}{R} \sum_{j=1}^{\infty} \left(\frac{R}{1+R} \right)^j \frac{1}{\bar{t}} \left(\frac{t}{\bar{t}} \right)^{jN-1} \frac{N^{jN}}{(jN-1)!} e^{-tN/\bar{t}} \qquad (10)$$

However, as shown in Figs. (**2c**) and (**2d**), the RTD curves for the co-current operation showed a bimodal behaviour whose peaks decreased with decreasing gas flow rate. As a result, the solid flow dynamics was modelled by a series of CSTRs in parallel with a plug flow reactor. The model also admitted interphase solid transport between the wake and bulk liquid phase and the solid RTD may be written:

$$E(t) = L^{-1} \left[\phi C_N(s) + (1-\phi) C_N^p(s) \right] \qquad (11)$$

where, $C_N(s)$ and $C_N^p(s)$ may be obtained as:

$$C_j(s) = \frac{\begin{bmatrix} M_1 C_{j-1}(s) + M_2 b C_{j-2}(s) e^{-2\bar{t}_p/N} \\ + \dots + M_2 b^{j-2} C_1(s) e^{-(j-1)\bar{t}_p/N} \\ + M_3 b^{j-1} C_o(s) e^{-j\bar{t}_p s/N} \end{bmatrix}}{\tau_m \left(s + a/\tau_m \right)} \qquad (12)$$

Similarly, unsteady-state mass balance around the PFR section yields,

$$C_N^p(s) = b^N C_o(s) e^{-t_p s} + M_4 b^{N-1} C_1(s) e^{-(N-1)t_p s/N} \qquad (13)$$
$$+ \dots + M_4 b C_{N-1}(s) e^{-t_p s/N} + M_4 C_N(s)$$

where, the variables M_1, M_2, M_3, M_4, a and b are functions of gas and liquid superficial velocities and are defined elsewhere [3].

For the same 50-RTD data points used in the neural nets evaluation, equations (10) and (11) gave a correlation coefficient value of 0.912 as seen in Table **2**. Although this indicates a good fit, the predictions are less accurate than most of the neural net configurations examined in this study. This illustrates the superiority of neural nets in the prediction of complex, nonlinear behaviour such as that implicated in RTD of solid particles in three-phase reactors.

Interpreting Neural Network Weights

The neural net weight matrix can be used to assess the relative importance of the various input variables on the output variables (in this case RTD of solids). On partitioning of neural net connection weights, the relative importance of jth input variable on n^{th} output variable may be expressed as [9, 13]:

$$I_{jn} = \frac{\sum_{m=1}^{m=N_h} \left(\left(|w_{jh}| \Big/ \sum_{k=1}^{N_i} |w_{ih}| \right) \times |w_{ho}| \right)}{\sum_{k=1}^{k=N_i} \left\{ \sum_{m=1}^{m=N_h} \left(\left(|w_{jh}| \Big/ \sum_{k=1}^{N_i} |w_{ih}| \right) \times |w_{ho}| \right) \right\}} \qquad (14)$$

The relative importance of various variables as calculated by equation (14) is shown in Table **3** and as expected, it is clear that the gas-liquid contacting scheme, with relative importance of 30 %, had the most marked influence on the RTD of solid particles. With a relative importance value of 20.4 %, the liquid velocity was found to be the second most important variable but closely followed the gas velocity and time with particle size being the least significant.

Table 3. Relative Importance of Variables Studied on the RTD of Solid Particles

Variable	Range	Relative importance
Operation	Co-current (+1) and Countercurrent (−1)	30.0 %
Gas velocity	0.05-0.5 cm s^{-1}	18.3 %
Liquid velocity	0.03-0.13 cm s^{-1}	20.4 %
Time	0-120 min	17.5 %
Particle size	0.2-200 micron	13.8 %

CONCLUSIONS

A neural net modeling of solid particles RTD in a three-phase reactor was carried out. A set of 300 input-output patterns was used to train a number of 3-layered ANN configurations having one hidden layer. The input layer of neurons consisted of 5 variables – gas and liquid velocities, particle size, time and contacting scheme. The hidden layer contained a variable number of neurons. The output layer had only one neuron – the RTD of solid particles. When tested with arbitrarily selected 50 data points, most of the configurations studied gave better results than the phenomenological model previously developed. However, a neural net with 1-hidden layer having 6-neurons was found to be most optimal giving a correlation coefficient value of 0.995 and mean absolute-percentage-error value of 3.43 %. A partitioning of the neural net weights was carried out to study the relative importance of the various input variables on the output variable (RTD. The relative direction of the gas and liquid flow (co- or counter-current mode) was found to be the most important variable for solids RTD in the bubble column reactor.

ACKNOWLEDGMENTS

NOMENCLATURE

C_p = Pearson correlation coefficient

e = error

$E(t)$ = residence time distribution (RTD) of solids, min^{-1}

E_a = actual value of RTD, min^{-1}

E_p = predicted value of RTD, min^{-1}

I = relative importance of a variable

N = number of tanks in the series of tank model

N_o = total number of neurons in the output layer

N_h = total number of neurons in the hidden layer

N_i = total number of neurons in the input layer

N_s = total number of input-output patterns

R = recycle ratio in the series of tanks model

$\bar{t}$ = average experimental residence time of solids, min

$\bar{t}_p$ = total residence time of solids in wake (PFR) in co-current RTD model, min

$T_{o,j}$ = target or actual prediction from the nth output neuron for j^{th} input-output pattern

w = weights of neural connections

x = input variable

y = output variable

$Y_{o,j}$ = neural net prediction from the nth output neuron for j^{th} input-output pattern

Subscripts

i = refers to input layer variables

h = refers to hidden layer variables

o = refers to output layer variables

Greek

α = momentum parameter (0.6–0.9)

β = learning rate parameter (0.6–0.9)

ϕ = fraction of solid particles which are transported *via* bulk liquid

τ_m = residence time of solids in one tank, min

REFERENCES

[1] M. Cassanello, F. Larachi, C. Guy, and J. Chaouki, "Solids mixing in gas-liquid fluidized beds: Experiments and modeling". *Chem. Eng. Sci.*, vol. 51, pp. 2011-2020, 1996.

[2] K. Kiared, F. Larachi, C. Guy and J. Chaouki, "Trajectory length and residence time distribution of the solids in three-phase fluidized beds". *Chem. Eng. Sci.*, vol. 52, pp. 3931-3939, 1997.

[3] V.K. Pareek, Z. Yap, M.P. Brungs and A.A. Adesina, "Particle residence time distribution in three-phase annular bubble column reactor". *Chem. Eng. Sci.*, vol. 56, 6063-6071, 2001.

[4] D.M. Himmelblau, "Application of artificial neural networks in chemical engineering". *Korean J. Chem. Eng.*, vol. 17, pp. 373-392, 2000.

[5] J. Michalopoulos, S. Papadokonstadakis, G. "Arampatzis and A. Lygeros, Modeling of an industrial fluid catalytic cracking unit using neural networks". *Trans. IChemE*, vol. 79, pp. 137-142, 2001.

[6] U. Rodemerck, M. Baerns, M. Holena and D. Wolf, "Application of a genetic algorithm and a neural network for the discovery and optimisation of new solid catalytic materials". *Appl. Surf. Sci.* vol. 223, pp.168-174, 2004.

[7] M.E. Gunay and R. Yildrim, "Neural network aided-design of Pt-Co-Ce/Al₂O₃ catalyst for selective CO oxidation in hydorgen-rich streams". *Chem. Eng. J.*, 2008, In press.

[8] E.J. Molga, "Neural network approach ot support modelling of chemical reactors: problems, resolutions, criteria of application". *Chem. Eng. Proc.*, vol. 43, 675-695, 2003.

[9] A. Elkamel, S. Abdul-Wahab, W. Bouhamra and E. Alper, "Measurement and prediction of ozone levels around a heavily industrialized area: A neural network approach". *Adv. Environ. Res.*, vol. 5, pp. 47-59, 2001.

[10] S.S. Tampe, B.D. Kulkarni and P.B. Deshpande, *Elements of Artificial Networks with Selected Applications in Chemical Engineering, and Chemical and Biological Sciences*, Louisville, KY, USA: Simulations and Advanced Control Ltd., 1996.

[11] D.E. Rumelhart, G. E. Hinton and R.J. Williams, "Learning Internal Representations by Error Propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* vol.1, D. E. Rumelhart, J. L. McClelland, Eds., Massachusetts: MIT Press, 1996, pp. 318-362.

[12] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representation by Back-Propagation Errors". *Nature*, vol. 323, pp. 533-536, 1986.

[13] G.D. Garson, "Interpreting Neural-Network Connection Weights". *AI Expert*, pp. 46-51, April 1991.