PREDICTION OF EXTRACTION EFFICIENCY IN RDC COLUMN USING ARTIFICIAL NEURAL NETWORK

Prof. Dr. Adil. A. A. Al-Hemiri and Chalak S. Omar

Chemical Engineering Department, College of Engineering, University of Baghdad

ABSTRACT

An application of neural network technique was introduced in modeling extraction efficiency in RDC column, based on a data bank of around 352 data points collected in the open literature. Three models were made, using back-propagation algorithm, the extraction efficiency was found to be a function of seven dimensionless groups: Weber number (we), (V_d / V_c), (μ_c / μ_d), (D_s / D_t), (D_r / D_t), (Z_c / D_t) and (Z_t / Z_c). Statistical analysis showed that the proposed models have an average absolute error (AARE) and standard deviation (SD) of 12.23% and 10.61% for the first model, 5.35% and 6.21% for the second model, 8.34% and 7.59% for the third model. The developed correlations also show better prediction over a wide range of operating conditions, physical properties and column geometry.

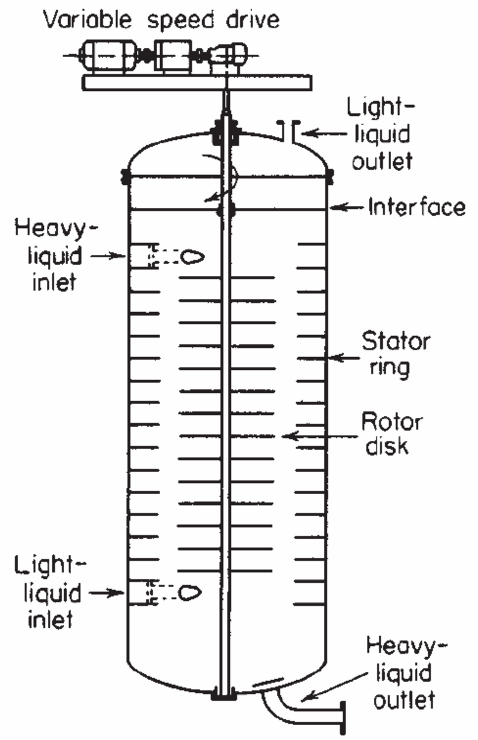
KEY WORDS

Rotating disc contactor, Extraction Efficiency, Artificial neural network, Back-propagation algorithm

INTRODUCTION

Liquid-liquid extraction has been emerging as a very important method for separation of liquid mixtures into its components by means of a solvent. The solvent used in the extraction process should be immiscible or partially miscible with one of the components of the mixture in order to facilitate the separation of the liquid phases (Laddha and Degaleesan, 1976).

The rotating disc contactor (RDC) has received considerable attention as liquid-liquid extraction equipment for refining of vegetable oils, processing of nuclear fuels, refining of crude petroleum and purification of vitamins. In common with other agitated columns it relies upon the application of mechanical energy to the contactor contents to gain high mass transfer efficiency (Laddha and Kannappan, 1978). It consists of a vertical cylindrical shell divided into a number of compartments by a series of stator rings. A rotating disc supported on a central shaft driven by an electric motor is located in each compartment. The feed inlets, at each end of the column, are arranged tangentially in the direction of rotation. The outlets are usually through the top and bottom plates of the column. The dense phase is introduced into the top of the column and the light phase into the bottom, so that, counter current flow is established by gravity. One of the phases is dispersed by the action of the rotating discs. At the ends of the column there are settling zones for phase separation (Alders, 1959), as shown in fig. 1. The advantages of such device are the acheivement high efficiency per unit height and high throughput for a given flow area. The column is relatively free from plugging, and thus can be operate in the presence of small amounts of suspended solids or other impurities. Moreover it requires low driving power and has comparatively low operation and maintenance costs (Zhang et al, 1981).



Rotating-disk (RDC) extractor. (*Courtesy of Glitsch Process Sys-*

COLUMN DESIGN VARIABLES:

The important column parameters that effect the performance of a rotating disc contactor for a given extraction system are;

Column diameter	(D_t)
Rotor disc diameter	(D_r)
Stator diameter	(D_s)
Compartment height	(Z_c)
Effective column height	(Z_t)
Speed of rotor disc	(N)

These dimensions are normally given as ratios with respect to (D_t). And may be varied slightly to provide flexibility in design. For the optimum design, the column dimensions should have the ratios (Remab and Olney, 1955; Logsdail et al, 1957).

$$(D_s/D_t) = 0.66 \text{ to } 0.75$$
$$(D_r/D_t) = 0.5 \text{ to } 0.66$$
$$(Z_c/D_t) = 0.33 \text{ to } 0.5$$

EXTRACTION EFFICIENCY:

Numerous studies have been made to obtain the effect of different parameters on RDC 's efficiency. Reman and Olney (1955) investigated the influence of column geometry and flow rates. Efficiency was found to increase with,

1. Decreasing stator opening.
2. Decreasing compartment height.
3. Increasing of dispersed flow rate at constant continuous flow rate.
4. Increasing rotor speed.
5. Increasing diameter of rotor discs.
6. Increasing specific load.

However, under certain conditions, increasing in rotor speed and specific load reduced the efficiency, due to back mixing.

Reman and Olney (1955) interpreted their results by plotting the efficiency, defined as the number of stages per foot column height, versus the energy input per unit volume ($N^3.R^5/H.D^2$). Data for two column diameters, 4 inches and 16 inches correlated well.

Later Logsdail et al (1957), using the system toluene-acetone-water and butylacetate-acetone-water, the water being the continuous phase throughput, showed that the overall values of the mass transfer coefficient or H.T.U. , could be correlated by the expression;

$$\left[\frac{(H.T.U.)_{oc}}{V_c}\left(\frac{g^2.\rho_c}{\mu_c}\right)^{1/3}\right].x = \left[\frac{x}{K_{oc}.a}\left(\frac{g^2.\rho_c}{\mu_c}\right)^{1/3}\right] = K\left(\frac{\mu_c.g}{\overline{V}_N^{\,3}.(1-x)^3.\rho_c}\right)^{2m/3}.\left(\frac{\Delta\rho}{\rho_c}\right)^{2(m-1)/3} \quad ..(1)$$

Use of this expression for design purposes necessitates evaluation of the constant K, the exponent m, and the characteristic velocity $\overline{V}_N$

This may be determined from tests with the given system in a small-scale laboratory column. Alternatively for the case of transfer from an aqueous into solvent phase, i.e., a case of hindered coalescence, $\overline{V}_N$ may be evaluated using the equation below;

$$\frac{\overline{V}_N.\mu_c}{\sigma} = 0.012\left(\frac{\Delta\rho}{\rho_c}\right)^{0.9}\left(\frac{g}{D_r.N^2}\right)^{1.0}\left(\frac{D_s}{D_r}\right)^{2.3}\left(\frac{Z_c}{D_r}\right)^{0.9}\left(\frac{D_r}{D_t}\right)^{2.7} \quad .....(2)$$

THE BACK-PROPAGATION ALGORITHM:

Back-propagation is a supervised learning technique used for training artificial neural networks. It was first described by Paul Werbos in 1974, and further developed by David E. Rumelhart, Geoffery E. Hinton and Ronald J. Williams in 1986.

It is most useful for feed-forward networks (networks that have no feedback, or simply, that have no connections that loop). The term is an abbreviation for "backwards propagation of errors". Back-propagation requires that the transfer function used by the artificial neurons (or "nodes") be differentiable. Back propagation networks are among the most popular and widely used neural networks because they are relatively simple and powerful.

. The input is the input to the hidden layer and the output layer is the output from the immediate previous layer, so it is called feed forward neural network. The number of the input units and the output units are fixed to a problem, but the choice of the number of the hidden units is somehow flexible as shown in fig. 2. Too many hidden units may cause over fitting, but if the number of hidden units is too small, the problem may not converge at all. Usually a large number of training cases may allow more hidden units if the problem requires so (Sivanadam, 2003).

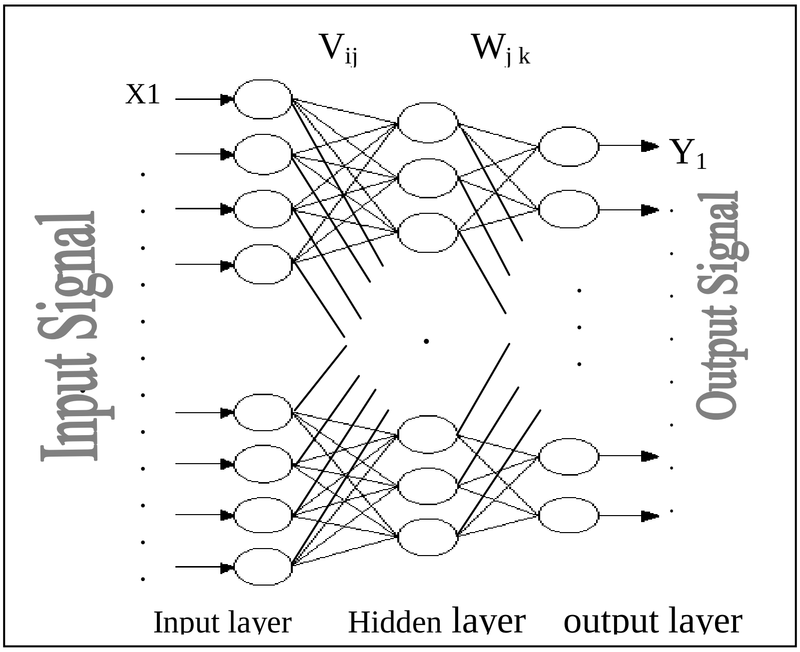


Fig. 2. Multi layer feed forward neural network

TRAINING A BACK-PROPAGATION NETWORK:

The conventional algorithm used for training a multi-layerd feed forward (MLFF) is the Bp algorithm, which is an iterative gradient algorithm designed to minimize the mean-squared error between the desired output and the actual output for a particular input to the network (Lendaris, 2004).

Two learning factors that significantly affect convergence speed as well as accomplish avoiding local minima, are the learning rate and momentum.

The learning rate (η) determines the portion of weight needed to be adjusted. However, the optimum value of η depends on the problem. Even though as small learning rate guarantees a true gradient descent, it slows down the network convergence process. If the chosen value of η is too large for the error value, the search path will oscillate about the ideal path and converges more slowly than a direct descent. The momentum (α) determines the fraction of the previous weight adjustment that is added to current weight adjustment. It accelerates the network convergence process. During the training process, the learning rate and the momentum are bring the network out of its local minima, and accelerate the convergence of the network.

The algorithm of the error back-propagation training for one hidden layer is given below (Lendaris, 2004):
Step1: initialize network weight values.

Step2: sum weighted input and apply suitable activation function to compute the output of the hidden layer.

$$h_j = f\left[\sum_i X_i W_{ij}\right] \qquad \ldots (3)$$

Where
h_j: The actual output of hidden neuron j for input signals X.
X_i: Input signal of input neuron (i).
W_{ij}: Synaptic weights between input neuron hidden neuron j and i.
f: The activation function.

Step3: sum weighted output of hidden layer and apply activation function to compute output of output layer.

$$O_k = f\left[\sum_j h_j W_{jk}\right] \qquad \ldots (4)$$

where
O_k: The actual output of output neuron k.
W_{jk}: Synaptic weight between hidden neuron j and output neuron k.

Step4: compute back propagation error.

$$\delta_k = (d_k - O_k) f'\left(\sum_j h_j W_{jk}\right) \qquad \ldots (5)$$

where
f' :The derivative of the activation function.
d_k: The desired output of neuron k.

Step5: calculate weight correction term.

$$\Delta W_{jk}(n) = \eta \delta_k h_j + \alpha \Delta W_{jk}(n-1) \qquad \dots (6)$$

Step6: sums delta input for each hidden unit and calculate error term.

$$\delta_j = \sum_k \delta_k W_{jk} \, f'\left(\sum_i X_i W_{ij} \right) \qquad \dots (7)$$

Step7: calculate weight correction term.

$$\Delta W_{ij}(n) = \eta \delta_j X_i + \alpha \Delta W_{ij}(n-1) \qquad \dots (8)$$

Step8: update weights.

$$W_{jk}(n+1) = W_{jk}(n) + \Delta W_{jk}(n) \qquad \dots (9)$$

$$W_{ij}(n+1) = W_{ij}(n) + \Delta W_{ij}(n) \qquad \dots (10)$$

Step9: repeat step2 for given number of error.

$$MSE = \frac{1}{2p}\left[\sum_p \sum_k (d_k^p - O_k^p)^2 \right] \qquad \dots (11)$$

Where p: The number of patterns in the training set.

Step10: END.

COMPUTER SIMULATION RESULTS:-

COLLECTION OF DATA:

In this work about 352 experimental points have been collected from 11 sources spanning the years 1954-1986. These data were divided into training part (75%) and testing part (25%). 224 data points were used in the first model which is for mass transfer from continuous to dispersed phase ($c \to d$), 128 data points used in the second model for mass transfer from dispersed to continuous phase ($d \to c$) and over all 352 data points used in the third model for mass transfer for both directions ($c \to d, d \to c$). These data includes six chemical systems with a large range of rotary speed, velocity of both continuous and dispersed phase, column geometry, also the physical properties for each chemical system. Table 1 gives the detailed listing of data used for the present work.

Table 1. Details of data bank used for the present work.

Author	System ….Dispersed phase ….Continuous phase	Direction of solute transfer	No. of data
Ismail (1985)	Toluene Acetone Water	$(c \rightarrow d)$	57
Al-Hemiri (1973)	Toluene Acetone water	$(c \rightarrow d)$ $(d \rightarrow c)$	46
Al-Hemiri (1973)	Liquid paraffin Methyle ethyl ketone water	$(c \rightarrow d)$ $(d \rightarrow c)$	22
Korchinsky (1982)	Toluene Acetone water	$(c \rightarrow d)$	15
Zhang(1985)	Kerosene n-butyric acid water	$(c \rightarrow d)$ $(d \rightarrow c)$	30
Al-Aswad (1985)	Clairsol 350 Acetone water	$(c \rightarrow d)$ $(d \rightarrow c)$	22
Al-Husseini (1986)	Toluene Acetone water	$(c \rightarrow d)$	52
Cruz-pinto (1983)	Toluene Acetone water	$(c \rightarrow d)$	4
Chartres and Korchinsky (1978)	Toluene Acetone water	$(c \rightarrow d)$	7
Korchinsky and Cruz-pinto (1979)	Toluene Acetone water	$(c \rightarrow d)$	4
Vermijs and Kramers (1954)	Water Acetic Acid Methyl isobutyl ketone	$(c \rightarrow d)$	38
Krishnaiah (1967)	Benzene Acetic acid water	$(d \rightarrow c)$	55

CHOICE OF INPUT PARAMETERS:

In this study there are thirteen parameters used, these are: rotor speed (N), dispersed phase superficial velocity (V_d), continuous phase superficial velocity (V_c), dispersed phase density (ρ_d), continuous phase density (ρ_c), dispersed phase viscosity (μ_d), continuous phase viscosity (μ_c), interfacial tension (σ), rotor disc diameter (D_r), stator diameter (D_s), column diameter (D_t), compartment height (Z_c) and column height (Z_t).

These parameters are input to the network as seven dimensionless groups that affect the efficiency of (RDC).

The main advantage of dimensionless group is to reduce the number of input parameters to the network.

The dimensionless group input to the network are:

1. Weber number: that consists the density difference of two phases ($\Delta\rho$), interfacial tension (σ), rotor disc diameter (D_r) and rotor speed (N)

$$We = \frac{\Delta\rho . D_r .(N . D_r)^2}{\sigma}$$

2. Ratio of dispersed to continuous phases velocities =(V_d / V_c) .
3. Viscosity ratio of the phases = (μ_c / μ_d).
4. Ratio of stator to column diameters = (D_s / D_t).
5. Ratio of disc to column diameters = (D_r / D_t).
6. Ratio of compartment height to column diameter = (Z_c / D_t).
7. Ratio of column to compartment height =(Z_t / Z_c).
   This is number of compartment.

THE STRUCTURE OF ARTIFICIAL NEURAL NETWORKS:

The ANN structure is determined by trial and error. For the first model it consists of seven input neurons in the input layer, twenty one neurons in the hidden layer and one neuron in the output layer, for second case it has the same structure but different in number of neurons in the hidden layer it consists of twenty four neurons. The last case has the same structure of previous models but different in number of neurons with nineteen neurons in the hidden layer. And then the networks trained with back-propagation algorithm then calculate the weights and biases matrices.

The trial and error to find the best ANN correlation model for the case ($c \rightarrow d$), case ($d \rightarrow c$) and for over all two cases together are shown in tables (2), (3) and (4) respectively.

Table (2) Some of the trial and error attempts for finding the best ANN model for the case (c $\rightarrow$ d).

Structure	MSE	No. of iteration	Learning rate	Momentum coefficient	Transfer function
[7-15-1]	0.01	2379	0.2	0.8	Tan sigmoid
[7-16-1]	0.005	1313	0.7	0.8	Tan sigmoid
[7-17-1]	0.008	1303	0.4	0.9	Tan sigmoid
[7-18-1]	0.003	1074	0.8	0.9	Tan sigmoid
[7-20-1]	0.002	2078	0.65	0.8	Tan sigmoid
[7-21-1]	0.001	2101	0.75	0.9	Tan sigmoid

Table (3) Some of the trial and error attempts for finding the best ANN model for mass transfer from ($d \rightarrow c$).

Structure	MSE	No.of iteration	Learning rate	Momentum coefficient	Transfer function
[7-15-1]	0.01	678	0.2	0.8	Tan sigmoid
[7-16-1]	0.008	721	0.5	0.8	Tan sigmoid
[7-17-1]	0.004	768	0.6	0.9	Tan sigmoid
[7-18-1]	0.003	362	0.7	0.9	Tan sigmoid
[7-24-1]	0.001	542	0.75	0.9	Tan sigmoid

Table (4) Some of the trial and error attempts for finding the best ANN model for mass transfer for all two cases**.**

structure	MSE	No.of iteration	Learning rate	Momentum coefficient	Transfer function
[7-15-1]	0.1	3180	0.5	0.9	Tan sigmoid
[7-16-1]	0.008	2352	0.6	0.9	Tan sigmoid
[7-17-1]	0.004	2718	0.7	0.8	Tan sigmoid
[7-19-1]	0.002	8320	0.75	0.9	Tan sigmoid

The weights and biases matrices for the three models are shown in equations below:

For the first model:

$$w_h = \begin{pmatrix} 0.6712 & -3.3673 & -1.9696 & 10.9379 & -15.2399 & 12.0679 & -0.5428 \\ -0.2061 & 0.0393 & -2.5898 & 5.9822 & 23.5954 & -4.8994 & 0.5979 \\ -0.0663 & -6.6292 & 2.4579 & 0.9583 & -39.1761 & 8.0865 & 0.2162 \\ -1.4236 & 9.1276 & -1.2101 & -1.0746 & -60.9264 & -0.1834 & 4.8962 \\ 0.2418 & -7.4610 & 11.5550 & -2.9587 & 23.0981 & -1.2386 & -0.4613 \\ 0.5020 & 6.7192 & -1.5953 & 5.5997 & -47.1539 & -1.1287 & -0.4449 \\ -0.1585 & 4.9900 & 2.8305 & -6.8283 & 35.0873 & 1.4067 & -0.5246 \\ 0.0117 & -6.6071 & 0.7065 & 1.2093 & -0.9628 & 6.0162 & -0.2302 \\ 0.0881 & -3.6348 & -0.5344 & -5.8485 & -37.5682 & -2.7069 & 0.0733 \\ -0.1074 & -0.4881 & 0.4455 & -9.2231 & -29.7618 & -2.9980 & 0.4807 \\ -0.0043 & -3.0572 & 1.4106 & -1.2227 & -31.5198 & -5.0645 & 0.0121 \\ 0.0049 & -1.3719 & 0.8255 & 15.8067 & -39.7171 & -8.2089 & -0.0138 \\ -3.1340 & -3.2529 & -0.1819 & -2.4573 & -30.3152 & 6.3427 & 0.1685 \\ -6.4694 & -20.1816 & 11.4064 & -4.6109 & 30.2074 & -1.0619 & 2.6669 \\ 0.0339 & 4.3099 & -3.1361 & 8.3781 & 12.1996 & 7.8483 & 0.0662 \\ -0.0617 & -7.1877 & 7.2925 & -3.4151 & 17.3276 & -6.9807 & -0.0961 \\ -0.0504 & 3.4739 & 3.5516 & 3.8803 & -15.4565 & -1.8829 & -0.0897 \\ 0.0011 & -6.8589 & 3.4766 & -10.6542 & 0.5341 & -4.1628 & 0.0705 \\ 0.1253 & 4.0017 & -3.8548 & 4.7763 & 38.7567 & -3.3763 & -0.1087 \\ 3.3161 & 8.4312 & -0.8098 & -3.1555 & 31.0113 & 8.4326 & -1.4653 \\ 0.0677 & -0.7855 & 0.8125 & 6.3748 & -8.6399 & -5.8645 & 0.2365 \end{pmatrix} \quad \ldots(12)$$

$$, \; b1 = \begin{pmatrix} 8.7672 \\ -23.9633 \\ 16.0111 \\ 33.8680 \\ 33.8680 \\ 24.9955 \\ -6.5297 \\ 1.0558 \\ 25.3744 \\ 13.1758 \\ 20.0840 \\ 12.9769 \\ 13.7197 \\ -17.5943 \\ -19.9880 \\ 3.2111 \\ 5.5152 \\ 9.1235 \\ -24.3990 \\ -17.2059 \\ 2.0656 \end{pmatrix} \quad \ldots(13)$$

$$b2 = [1.1006] \quad \ldots(14)$$

$$w_0 = [-3.4891 \quad -8.9789 \quad 6.1570 \quad -1.3785 \quad -2.6963 \quad 2.1895 \quad 10.3487 \quad -5.3463 \quad 10.0372 \quad 7.6532 \quad -11.0769 \quad 6.0423 \quad 6.2067 \quad -17.6195 \quad -8.3889 \quad -4.9637 \quad 1.2610 \quad 3.6164 \quad 7.3337 \quad -19.1532 \quad 1.7628] \quad \text{....(15)}$$

For the second model:

$$w_h = \begin{pmatrix}
0.1234 & 0.2889 & -0.2098 & 3.4187 & 35.7059 & 2.6383 & -0.1933 \\
-0.0357 & 0.7395 & -0.0531 & 4.1740 & -13.3786 & 3.6104 & 0.3944 \\
0.0930 & -0.3501 & 0.0753 & 10.3319 & -23.1379 & 3.9859 & 0.0543 \\
-0.6547 & 0.0229 & 0.2226 & -3.9909 & 1.1249 & 5.5466 & 0.1787 \\
0.0956 & -0.0465 & 0.0034 & -1.7553 & -60.4389 & 0.2748 & -0.0457 \\
-0.0011 & -1.4635 & 1.5229 & -0.6160 & 4.4112 & 5.8333 & -0.1594 \\
0.1889 & -0.8623 & -0.1330 & 4.4915 & 35.6777 & 0.5701 & 0.0534 \\
-0.0900 & 0.5578 & -0.0051 & -7.9568 & -72.2110 & 2.3670 & 0.0189 \\
0.1622 & 0.4424 & 0.0356 & -10.2791 & 42.3113 & 8.6717 & 0.1048 \\
-0.0716 & -0.1468 & -0.0947 & 6.5884 & 41.9295 & 6.5861 & -0.2150 \\
0.0011 & -0.7969 & 0.1646 & 0.4433 & 37.0766 & -5.1581 & 0.1260 \\
-0.4757 & -0.4012 & -0.0098 & -9.3298 & -7.5400 & 2.5654 & -0.3187 \\
0.1050 & 0.7325 & -0.1837 & -8.4602 & -54.0223 & -7.0581 & -0.1633 \\
0.0316 & -0.5672 & -0.3076 & -7.5982 & 43.0925 & -4.9177 & 0.0049 \\
-0.2175 & -0.3830 & 0.0075 & -2.4090 & -48.6293 & 7.1045 & 0.0400 \\
-0.1616 & 0.6706 & -0.0360 & -2.3382 & -8.4124 & 3.8209 & -0.1280 \\
-0.0489 & 0.1188 & -0.1397 & -8.7265 & -57.7659 & -8.5403 & -0.1149 \\
-0.1628 & -0.2300 & -0.1223 & 12.8269 & -33.7744 & 1.6003 & -0.0373 \\
0.0281 & 0.6424 & -0.1755 & -10.5666 & 68.4840 & -1.2927 & 0.0101 \\
0.5449 & 0.0258 & 0.0105 & -2.9365 & 51.8078 & 9.2572 & -0.1245 \\
-0.1760 & -0.5069 & -0.3172 & -5.9202 & -31.6383 & -6.1096 & 0.1793 \\
0.4355 & -0.5017 & 0.0074 & -6.3292 & 44.2235 & -7.4766 & 0.1543 \\
-0.1672 & 0.0926 & 0.4492 & 10.6306 & -3.2991 & -0.5583 & 0.6760 \\
-0.0400 & -1.8825 & 0.2747 & -4.7801 & 10.1719 & 11.9353 & -0.0195
\end{pmatrix} \text{....(16)} \quad , \quad b1 = \begin{pmatrix}
-22.4296 \\ -3.2096 \\ 6.9129 \\ -2.1792 \\ 37.2682 \\ -0.5794 \\ -17.6551 \\ 43.7335 \\ -18.8823 \\ -28.9934 \\ -19.2918 \\ 11.2014 \\ 37.5087 \\ -18.3603 \\ 23.8124 \\ 0.3039 \\ 36.8954 \\ 10.3930 \\ -28.2146 \\ -29.5656 \\ 21.5242 \\ -14.6105 \\ -7.9138 \\ -2.0174
\end{pmatrix} \text{....(17)}$$

$$b2 = [0.0214] \quad \text{.....(18)}$$

$$w_o = [0.6345 \quad 0.4762 \quad -0.5996 \quad 0.1780 \quad 0.3690 \quad 1.3332 \quad 1.1331 \quad 1.0184 \quad 0.1551 \quad 0.7346 \quad 1.6231 \quad -0.0963 \quad 0.9227 \quad 0.6875 \quad 0.0962 \quad 0.2680 \quad -0.6648 \quad -0.6382 \quad -0.9553 \quad -0.8878 \quad -0.6872 \quad -0.0673 \quad 0.4206 \quad -0.7015] \quad \text{......(19)}$$

For the third model:

$$w_h = \begin{pmatrix}
0.0249 & 1.3648 & -0.4491 & 6.3232 & 14.8117 & 3.6421 & 0.1036 \\
-0.5099 & 6.1151 & -3.4781 & -2.5083 & -29.4055 & 8.2767 & 1.0769 \\
-0.2348 & 0.1972 & 0.1672 & 6.2039 & 33.9799 & -1.8162 & 0.1033 \\
0.0389 & 0.1031 & -0.2641 & 6.1485 & 32.9483 & 6.0856 & -0.2337 \\
-0.0083 & 7.5839 & -0.1512 & 7.5159 & -24.7785 & -7.7478 & 0.2997 \\
0.6398 & -2.2457 & -1.4695 & 6.8300 & 13.4595 & 3.5864 & -22.9526 \\
-0.1163 & -0.9637 & -0.0369 & -7.0662 & 23.4879 & -5.2558 & 0.0320 \\
-0.0034 & 11.6483 & -1.5612 & -9.4009 & 17.8948 & 2.6496 & -0.0233 \\
-0.0057 & -3.1547 & 0.0702 & -8.9703 & -5.5169 & 12.0045 & -0.5636 \\
0.0155 & 1.8149 & 2.1384 & 3.6038 & -3.8340 & 3.4886 & -0.9107 \\
0.4714 & 2.1124 & -3.7127 & -8.5942 & -4.6193 & 7.6501 & -1.0597 \\
-0.0031 & -1.1453 & 0.7689 & -13.9756 & 24.5252 & 4.0808 & 0.0316 \\
-0.0182 & -0.2930 & -5.5158 & 10.5766 & -1.3630 & 8.9686 & -0.0562 \\
0.1329 & -5.2441 & -0.0645 & -11.1016 & -13.7026 & 8.0867 & -1.0243 \\
0.1777 & 4.4825 & -0.6572 & 0.9513 & 2.4935 & 3.7738 & -0.5070 \\
-0.0015 & 0.0619 & 0.8831 & -3.2606 & 2.1586 & -7.8348 & 0.0943 \\
-0.0022 & -0.0247 & 3.2039 & 12.9280 & -31.2622 & -0.6524 & -0.2862 \\
-0.1088 & -1.7109 & 3.5722 & 10.5591 & 18.5896 & -5.4965 & 0.0918 \\
-0.9543 & 0.1094 & 0.0729 & -10.0114 & -5.1432 & -6.9029 & -0.0312
\end{pmatrix} \quad , \quad b1 = \begin{pmatrix}
-18.6431 \\ 17.0586 \\ -21.0471 \\ -18.3708 \\ 7.2773 \\ -15.4990 \\ -4.3358 \\ -5.8046 \\ 17.4021 \\ 8.6129 \\ -0.1512 \\ -4.3949 \\ -4.2431 \\ 9.2810 \\ 3.6552 \\ 1.8434 \\ 7.8422 \\ -20.6339 \\ 14.1979
\end{pmatrix}$$

…..(20) …(21)

$b2 = [-1.4308]$ …..(22)

$w_o = $[1.5043 0.3633 -0.8227 -1.9027 -13.4247 -18.5098 1.8115 -0.5788
-7.8875 -0.8821 1.2929 0.9230 -5.8703 -4.1934 0.4554
-11.1639 7.4440 -1.5360 5.0601
………(23)

TEST OF THE PROPOSED ANNS:

The ANN models were tested using another data set to show the accuracy of the network for prediction extraction efficiency in (RDC). The first model was used to generate (60) new data values for mass transfer from (c$\rightarrow$d), the second model was used to generate (38) new data values for mass transfer from ($d \rightarrow c$) and the third model was used to generate (98) new data values for mass transfer from ($c \rightarrow d, d \rightarrow c$). The comparison between experimental and predicted efficiency for three cases were plotted in figures below:

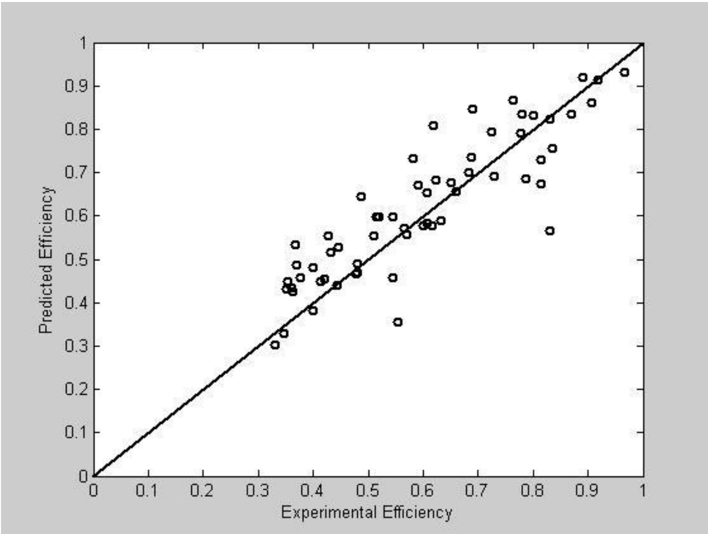


Fig. 3 .Comparison between experimental and predicted efficiency for the case ($c \rightarrow d$) in testing set.

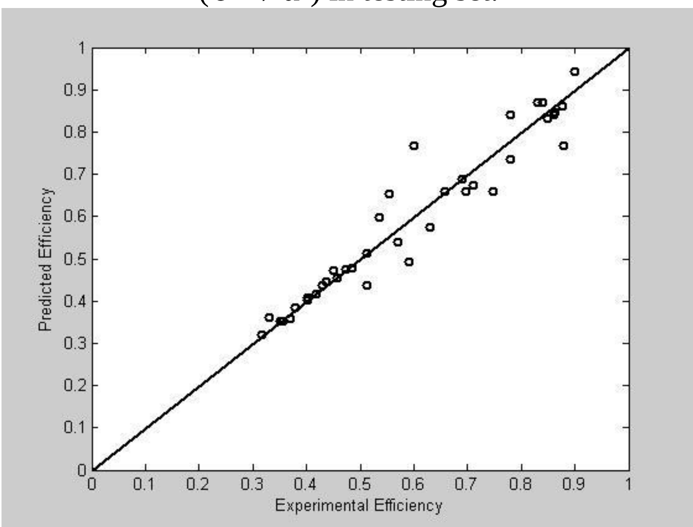


Fig. 4 . Comparison between experimental and predicted efficiency for mass transfer from ($d \rightarrow c$) in testing set.

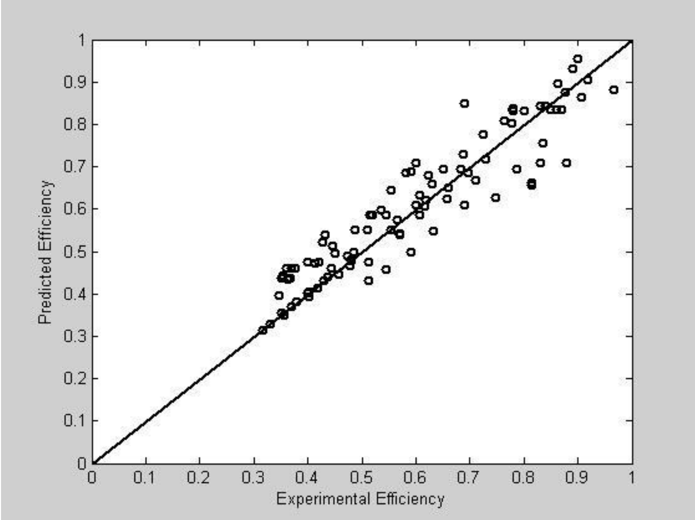


Fig. 5 . Comparison between experimental and predicted efficiency for all cases in testing set.

DISCUSSION

1- In the present work an attempt was made to correlate an ANN model for extraction efficiency prediction in (RDC). Three models were proposed, a model for mass transfer from ($c \rightarrow d$), mass transfer from ($d \rightarrow c$), and a model for the two cases together ($c \rightarrow d, d \rightarrow c$).

The accuracy of each model was validated by testing it with an experimental data not used in the training set and then compares the efficiency predicted from the ANN model with the experimental efficiency. Figures (3), (4) and (5) show the comparison between the predicted and experimental efficiency for ($c \rightarrow d$), ($d \rightarrow c$) and for ($c \rightarrow d, d \rightarrow c$) respectively.

Also the accuracy of these models was validated by statistical analysis (AARE, S.D and R). The model gives best output prediction based on AARE and S.D values respectively as shown in table 5.

Table 5. Statistical analysis information of three neural networks models

ANN models	Structure	AARE%	S.D%	R
Case 1 ($c \rightarrow d$)	[7-21-1]	12.23	10.61	0.879
Case 2 ($d \rightarrow c$)	[7-24-1]	5.35	6.21	0.962
Both cases	[7-19-1]	8.34	7.59	0.938

2. In this work back-propagation algorithm was used. This algorithm uses the (trainlm) training function, which appears to be the fastest method for training feed forward neural network.

3. The tansig (Hyperbolic tangent sigmoid) transfer function was used in the neurons in the hidden layer may be more accurate and is recommended for applications that require the hyperbolic tangent, especially prediction. Because the output from tan sigmoid varying from -1 to +1. The output neuron has a log sigmoid transfer function "logsig". Which is the best transfer function used in the output neuron for efficiency prediction. Because it generates outputs between 0 and 1.

4. Neural networks often encounter the well known 'overfitting' problem, which can make use of the ANN unreliable. To avoid 'overfitting' and make the ANN more useful, the following approach was used. The whole database was split into two parts, learning and generalization. The first part, called the 'learning file', was used to perform minimization using the ANN. The remaining part, called the 'generalization file', was used to validate the model

5. Third model concerning the two cases combined which cover cases of mass transfer from ($c \rightarrow d$) and from ($d \rightarrow c$) is found to be flexible and more comprehensive. Also it has optimum structure.

CONCLUSIONS:

It is very difficult to know which training algorithm will be the fastest for a given problem. It will depend on many factors, including the number of the experimental data points in the training set, the desired output (target) from the network, the relationship between the input and the desired output, the complexity of problem and the error goal. The number of neurons in the hidden layer was arrived at by trial and error starting from a minimum of fifteen neurons and according to the Hecht number.

ANN model can predict the extraction efficiency for a wide range of physical properties, operating parameters and column geometry. It has been demonstrated that the optimal model is a network that predicts for the two cases together ($c \rightarrow d, d \rightarrow c$) with one hidden layer.

NOMENCLATURE

a = Interfacial mass transfer area, m^2/m^3.

b = Bias.

d_k = The desired of output neuron k

D_r =Diameter of rotary disc, m.

D_s = Stator ring diameter, m.

D_t =Diameter of RDC column, m.

f = The activation function.

f' =The derivation of the activation function.

g = Acceleration due to gravity, m/s^2.

h_j = The actual output of hidden neuron j.

K_{oc} =Over all Mass transfer coefficient, m/s.

K= Constant (dimensionless).

N= Speed of rotor disc, rps.

o_k = The actual output of output neuron k.

P =The number of patterns in the training set.

R =Correlation coefficient.

V_c = Velocity of continuous phase, m/s.

V_d = Velocity of dispersed phase, m/s.

$\bar{V}_N$ = Characteristic velocity, m/s.
W_{ij}, w_h = Synaptic weights between input and hidden neurons.
W_{jk}, w_0 = Synaptic weight between hidden and output neuron.
We = Weber number (dimensionless).
x = Hold up.
x_i = Input signal of input neuron i.
Z_c = High of compartment, m.
Z_t = High of RDC column, m.

GREEK SYMBOLS

α = Momentum rate.
δ_k =The error term
η = The learning rate
ρ =Density, kg/m^3.
μ =Viscosity, kg/m.s.
σ = Interfacial tension, N/m.
$\Delta\rho$ =Difference in density, kg/m^3.

SUBSCRIPTS

c continuous phase.
d dispersed phase.

ABBREVIATIONS

AARE= Average Absolute Relative Error.
BP=Back Propagation.
H.T.U =Height Transfer Unit, m.
MLFF= Multi-Layer Feed Forward.
MSE=Mean Square Error.
S.D=Standard Deviation.

REFERENCES:

- Al-Aswad, K.K., Mumford, C.J. and Jeffreys, G.V., (1985), “The application of drop size distribution and discrete drop mass transfer models to assess the performance of a rotating disc contactor ”, AICHE. J., 31, 9, 1488-1497.
- Alders, L., (1959), “liquid-liquid extraction”, Elsevier publishing company.
- Al-Hemiri, A.A., (1973), “The effect of surface renewal on mass transfer in agitated contactors”, Ph.D., thesis, University of Aston, UK.
- Al-Husseini, R. and Korchinsky, W.J., (1986), “liquid-liquid extraction (RDC) model parameters from drop size distribution and solute concentration measurements”, J. Chem. Tech. Biotechnol, 36, 395-409.
- Charteres, R.H., and Korchinsky, W.J., (1978), “Drop size and extraction efficiency measurements in a pilot plant rotating disc contactor, Trans. I. Chem. Eng, 56, 91-95.
- Cruz-Pinto, J.J.C., and Korchinsky, W.J., (1983), “Exact solutions of the Newman, and the Handlos-Baron, model equation for counter current flow extraction”, Comp. Chem. Eng, 7, 19-25.

- Ismail, A., (1985), "Evaluation of the hydrodynamic and mass transfer parameters in an extraction column (RDC)", Ph. D. thesis, Victoria, University of Manchester.
- Korchinsky, W.J., and Cruz-Pinto, J.C., (1979), "Mass transfer coefficients-calculation for rigid and oscillating drops in extraction columns", Chem. Eng. Sci, 34, 551-561.
- Korchinsky, W.J., Loke, C.T. and Cruz-pinto, J.J.C., (1982), "Optimization of drop size in counter current flow liquid-liquid extraction columns", Chem. Eng. Sci., 37, 5, 781-786.
- Krishnaiah, M.M., Pai, M.U., Rao, R. and Sastri, S.R., (1967), "Performance of a rotating disc contactor with perforated rotors", British chemical engineering, 12, 5, 719-721.
- Laddha, G.S., and Degaleesan, T.E., (1976), "Transport phenomena in liquid extraction", Tata Mc Graw-Hill publishing Co. Ldt, New Delhi.
- Laddha, G.S., and Kannappan, R., (1978), "Hydrodynamic and mass transport in rotating disc contactors", Can. J. Chem. Eng, 56, 137-149.
- Lendaris, G., (2004), "Supervised learning in ANN from introduction to artificial intelligence", New York, April 7.
- Logsdail, D.H., Thornton, J.D., and Pratt, H.R.C., (1957), "liquid-liquid extraction part XII: flooding rates and performance data for a rotating disc contactor", Trans. Inst. Chem. Eng., 35, 301-315.
- Reman, G.H, (1952), Proc.3rd world petrol congress, Hague, section III.
- Reman, G.H., and Olney, R.B., (1955), "The rotating-disc contactor- a new toot for liquid-liquid extraction", Chemical Engineering progress, 51, 3, 141-146.
- Sivanadam, S.N., (2003), "Introduction to artificial neural networks", Vikas publishing House Pvt. Ldt.
- Vermijs, H.J., and Kramers, H., (1954), "Liquid-liquid extraction in rotating disc contactor", Chem. Eng. Sci., 3, 55-64.
- Zhang, S.H., NI, X.D. and Su, Y.F., (1981), "Hydrodynamics axial mixing and mass transfer in rotating disc contactors", Can. J. Chem. Eng, 59, 573-583.
- Zhang, X.D., and Su, Y.F., (1985), "A model for liquid-liquid extraction column performance the influence of drop size distribution on extraction efficiency", Can. J. Chem. Eng, 63, 212-226.