ARTIFICIAL NEURAL NETWORK FOR PACKED COLUMN

Sandeep Palve[1], Nilesh Kshirsagar[2], S.M. Walke[3]*

[1,2,3] Department of Chemical Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai, India

* Corresponding Author: Email: wsantosh@yahoo.com Tel: +919594995879

Abstract— The generalization performances of the Back Propagation Multi-Layer Perceptron (BPMLP) and the Radial Basis Function (RBF) neural networks were compared together by resorting to several sets of experimental data collected from a pilot scale packed absorption column. The experimental data were obtained from an 11 cm diameter packed tower filled with 1.8 m ¼ inch ceramic Rashig rings. The column was used for separation of carbon dioxide from air using various concentrations and flow rates of Di-Ethanol Amine (DEA) and Methyl Di-Ethanol Amine (MDEA) solutions. Two in-house efficient algorithms were employed for optimal training of both neural networks. The simulation results indicated that the RBF networks can perform more adequately than the MLP networks for filtering the noise (measurement errors) and capturing the true underlying trend which is essential for a reliable generalization performance.

Index Terms— Multi-Layer Perceptron; Packed absorption column; neural networks; Rashig rings

IJSER

1 INTRODUCTION

THE Separation of carbon dioxide from air and various industrial gases is essential from both operational and environmental views. For example, CO2 must be separated from natural gas to increase its heating value or carbon dioxide is usually extracted from various flue gases in beverage industry. To reduce global warming, CO2 should be also removed from industrial flue gases before exhausting them to atmosphere. In many practical applications, the natural gas contains around five percent carbon dioxide (e.g. Iranian Mozduran sour natural gas contains 6.5% CO2). On the other hand, flue gases resulting from complete combustion of almost pure natural gases (e.g. Iranian Mozduran sweetened natural gas contains 98.4% CH4) usually contain around 9% CO2 when natural gas is burnt with 5% or more excess air. The experimental CO2 concentrations (1&5%) were selected to cover approximately the above specifications while meeting the CO2 flow-meter restrictions. Although several adsorption and membrane processes are recently used for CO2 separation purposes absorption processes are still more popular in this area Alkanolamines (such as DEA or MDEA) are usually used forefficient separation of carbon dioxide from various industrial gases. Packed towers usually provide higher mass transfer areas and lower pressure drops when compared to tray towers which are investigated the

* *Mr. Sandeep Palve is currently persuing her Bachelors Degree in Chemical Engineering at Mumbai University, India*
Email:sandy.palve@gmail.com
Mr. Nilesh Kshirsagar who is a co-author is persuing her Bachelors Degree in Chemical Engineering at Mumbai University, India
Email: nilesh_the_billa@yahoo.co.in

absorption of carbon dioxide from nitrogen using MEA and MDEA solutions in a packed column under various operating conditions. A two parameter theoretical model was presented for describing the CO2 absorption behavior. The proposed model was validated using test data. They concluded that "an increase in the absorption load due to increased inlet CO2 concentration or gas flow rate leads to a much shorter breakthrough time. However, an increase in the amine concentration significantly enhances the CO2 absorption". presented a theoretical model to investigate the effect of various operational parameters on the performance of a regeneration packed column. The experimental data were then correlated to estimate the water evaporation rate from desiccant (CaCl2) at various operating conditions. They concluded that "the water evaporation rate increases with increase of air and solution inlet parameters, namely, flow rate and temperature". used a non-equilibrium heat and mass transfer model to describe the chemical absorption of ammonia, carbon dioxide and hydrogen sulfide in an aqueous solution containing sodium hydroxide, MEA and MDEA. The chemical reaction and its influence on mass transfer in the electrolyte system were accounted by enhancement factors. The calculation of the hydraulic parameters was based on standard correlations. The predictions of the mass transfer model were validated using experimental data. employed back propagation artificial neural networks to investigate the fault diagnosis in an ammoniaewater packed distillation column. The network was reported to perform satisfactorily on detection of the designated faults. The relative importance of various input variables on the output parameters was calculated by partitioning the connecting weights. The simulation results indicated that "bottoms temperature, overhead composition and overhead temperature are not much affected by the disturbances in feed rate, feed composition and vapor rate in the given range". A complex computational mass transfer model (CMT) for modeling of chemical absorption processes in packed columns. The model was able to consider heat effect for prediction of the concentration and temperature profiles as well as the velocity distributions. The presented model coupled the computational fluid dynamics (CFD) with computational heat transfer (CHT). The model was successfully validated using borrowed experimental data collected from a 0.1 m ID and 7 m height pilot scale tower randomly packed with ½ inch ceramic Berl saddles. The column was used for chemical absorption of CO2 from air by aqueous monoethanolamine (MEA) solution at total pressure of 103.15 kpa. Other sets of literature data collected from an industrial-scale packed column (1.9 m ID and 26.6 m height) randomly filled with 2 inch stainless steel Pall rings were also used for validation purposes. Chemical absorption of CO2 from natural gas was conducted using aqueous MEA solution.

They argued that "the common viewpoint of assuming constant turbulent mass transfer diffusivity (Dt) throughout the entire column is questionable, even for the small size packed columns", since Dt varies along both axial and radial directions. In this article, the generalization performances of Multi Layer Percep-

IJSER

tron (MLP) and Radial Basis Function (RBF) artificial neural networks are compared together using a pilot-scale packed absorption column. The training data were collected by conducting several experiments on absorption of CO2 from air using various DEA and MDEA solutions with different concentrations. Two different points will be emphasized in this article: a) Superior performance of RBF networks (when equipped with proper regularization level) to MLP networks and conventional software's for empirical modeling of the absorption process. b) Providing experimental data for absorption of CO2 using a packed tower, which is quite limited and is essential for better analysis of the entire process.

2. OVERVIEW OF NEURAL NETWORKS :-

Neural networks generally consist of several interconnectedneurons in one or more hidden layers. They can be classified from different points of views such as the type of input transformation, their structural architecture and the type of learning algorithm. Neural networks may employ either projection or kernel based transformations to account for correlation among the inputs. In the first transformation the inputs are projected on a single axis, the projection may be linear or non-linear. The McCulloch-Pitt neuron, Perceptron and Adaline are examples of linear projections .In the second transformation, the norm (usually Euclidean) of the input vector with respect to a fixed point (center) is used.

Radial basis function networks are the most popular examples of the kernel based input transformations Neural networks are commonly classified based on their direction of signal flow into feed-forward and recurrent networks. In a feed-forward network, signals flow from the input layer to hidden layer(s) and then to the output layer via unidirectional connections. The neurons are connected from one layer to the next but not within the same layer. These networks can most naturally perform static mappings between input and output spaces. In other words, the output of a feed-forward network at a given instant is only a function of its input at the same instant. In general, two different classes of single and multiple layer feed-forward network architectures can be identified. In its simplest form, a feed-forward network is constructed from an input layer of source nodes that are projected onto an output layer of computation nodes via synaptic weights. The "single-layer" designation refers to the output layer containing feed-forward computation nodes (neurons). A linear associative memory is an example of a single-layer neural network. The network associates an output pattern (vector) with an input pattern (vector), and the information is stored in the network by virtue of the modifications made to the synaptic weights of the network.

Multi-layer feed-forward neural networks contain one or more hidden layer(s), whose corresponding computational nodes are called hidden neurons or hidden units. The function of the hidden neurons is to intervene between the external input and the network response.By adding one or more hidden layers, the network is enabled toextract higher order statistics (more information) by virtue of theextra set of

IJSER

synaptic connections and increased neural interactions.The ability of the hidden neurons to extract higher order statistics iscrucial for large input dimensions. The neurons of each layer may beeither partially or fully connected to the neighboring layers.In a recurrent network, the outputs of some neurons are feedbackto the neurons of the same layer or to the nodes of thepreceding layers. The signals can therefore flow in both forwardand backward directions. Recurrent networks have dynamic memories, that is their outputs at a given instant reflect both the current input as well as past inputs and outputs.

Due to their dynamic memory, recurrent neural networks are particularly suited for control applications and dynamic simulations. The learning algorithm of a neural network deals with the adjustment of the network parameters and usually settles to solving an unconstrained or constrained optimization problem. The model representing the neural network may be linear, non-linear or a combination of both with respect to the network parameters. The merit function characterizing the network performance may depend on the inputs alone or both the inputs and the outputs. The former objective function leads to learning without a teacher (unsupervised learning) while the latter objective requires a teacher to direct the learning (supervised learning). The learning process of an unsupervised network does not require target (measured) output(s).

Only input patterns are presented to the network during training. For a fixed architecture, the network automatically adapts the network parameters to cluster the input pattern into groups of similar features.In many engineering applications we are concerned with the estimation of an underlying trend (or function) from a limited number of input output data points with little or no knowledge of the form of the true function (truth). This problem is sometimes referred to as non-parametric regression, function approximation, system identification or inductive learning. In neural network parlance, it is usually called supervised learning. The underlying function is learned from the exemplars which a teacher supplies. The set of examples (the training set) contains elements that consist of paired values of the independent (inputs) and the dependent (outputs) variables. A supervised learning algorithm adjusts the network parameters according to the differences between the measured response $y(x_i)$ and the network outputs $\hat{y}(x_i)$ corresponding to a given input xi. Supervised learning requires a supervisor, to provide the target signals.

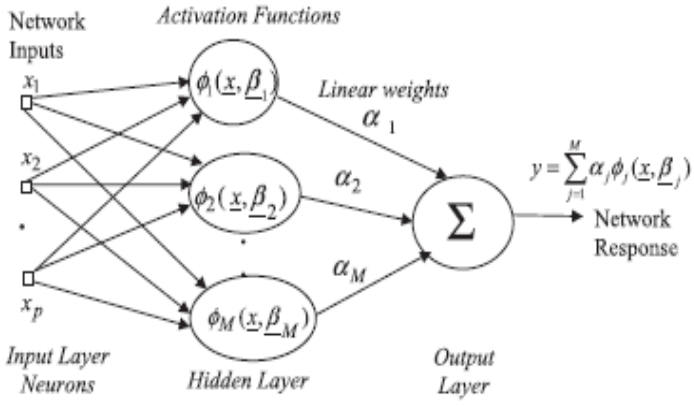


Fig.1.Schematic representation of a feed-forward neural network.

IJSER

3. EXPERIMENTAL SETUP

A pilot scale packed column was used to collect the required experimental data. Fig. 3.1 represents the schematic diagram of the packed tower and the corresponding auxiliary equipments. Two separate glass columns (ID ¼ 4.5 inches), one mounted over the other, and each one of them packed with 90 cm of¼ inches ceramic Rashig rings were used. A liquid re-distributor was assembled between the two packed sections. Sampling points were fitted at both ends of the column for collection of pressure drop data and concentration analysis purposes. An 80 L stainless steel solvent storage tank was used. The solvent and air rates were measured with separate glass rotameters prior to entering the tower. Large cylinders were used to provide the required carbon dioxide. The CO2 streamwas also measured via a small rotameter before mixing with air and entering the tower. The Hempl apparatus (as shown in Fig.3.) was used to measure the carbon dioxide concentrations at inlet and outlet air streams. These measured concentrations were occasionally verified (and corrected) by titrating both inlet and outlet solvent streams and using the component mass balance around the packed sections

Experimental data:-

After calibration of the Hempl gas analyzer apparatus, various measurements of inlet and outlet gas concentrations were performed. The following operating variables were varied during the experiments: U Type and solvent (DEA, MDEA and pure water), U Gas and liquid flow rates, U Concentrations of both solvents and gas streams. The small temperature fluctuations during each experiment were ignored and the average temperaturewas calculated using the initial and final conditions. The barometric pressure was close toone standard atmosphere for all experiments. Figs. represent the collected data for various operating conditions as specified in the corresponding figures. The error bars were computed from the following equations using multiple measurements at selected points

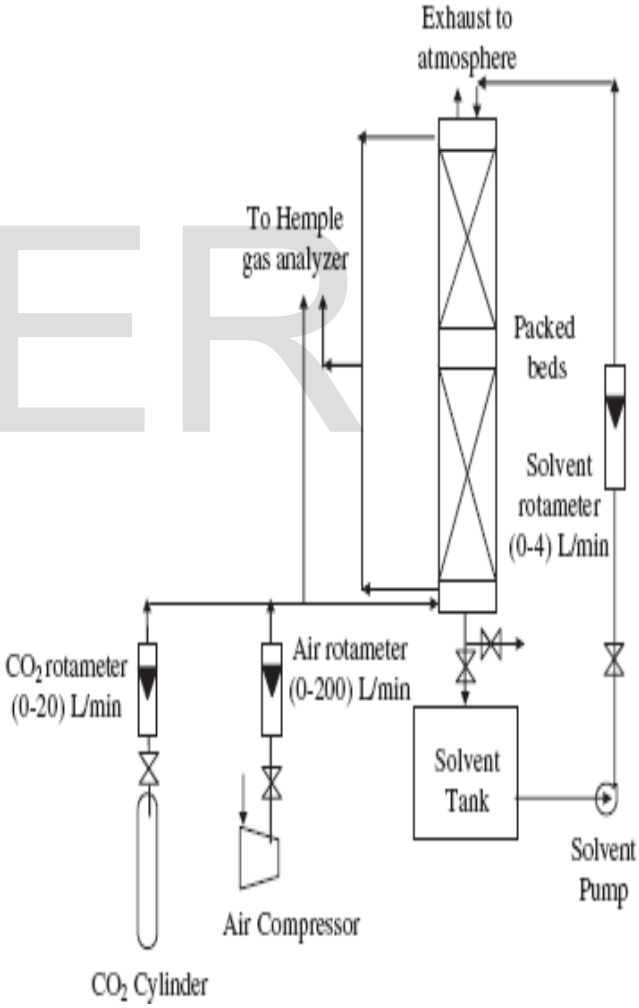


Fig 2. Schematic diagram of the pilot scale packed absorption system

IJSER

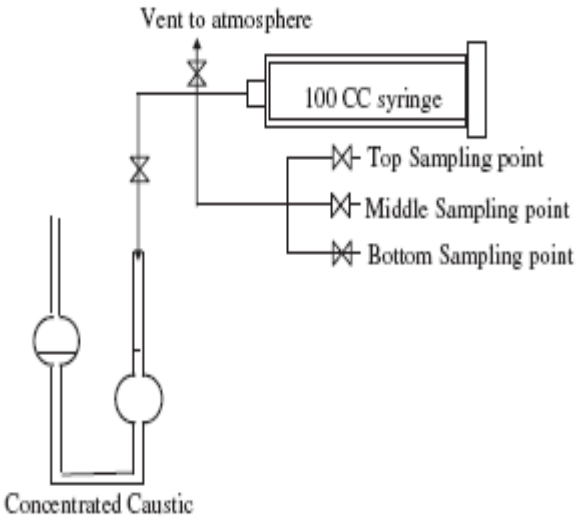


Fig 3. Schematic diagram of the Hempl gas analyzer.

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$\sigma^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1}$$

For small CO2 concentrations of entering air streams and assuming constant total gas flow rate across the entire column, the percent CO2 absorbed from air was computed for both figures as following:

$$\%(CO_2)\ \text{absorbed} = \frac{y_{inlet} - y_{outlet}}{y_{inlet}} * 100$$

Evidently, the outlet gas concentration increases by increasing the inlet CO2 mole fraction at constant solvent concentration as shown in Fig. 3. On the other hand, increasing the CO2 mole fraction of entering air (provided that all other conditions are kept constant) although increases the absorption rate (due to relatively larger driving forces), however it may not necessarily lead to higher percents of CO2 absorption.

Actually, in most cases, the increase in mass transfer rate is much smaller than the increase in inletconcentration due to limitations of equilibrium concentrations (or limited solvent capacity). In such cases, the percent of CO2 absorption will decrease with increase in the CO2 mole fraction of entering air. Evidently, the absorption capacity of the solvents increases by increasing the DEA or MDEA concentrations of the enteringsolutions. Furthermore, since DEA is much stronger alkali compared to MDEA, hence, DEA solutions are able to absorb morecarbon dioxide than MDEA solutions (at constant solventconcentrations).

4. HYSYS AND ASPEN SIMULATION RESULTS

The pilot scale process illustrated in Fig. 1 was simulated using HYSYS and Aspen technical software. Various built in thermodynamic packages were tried during the simulations. Fig. 5 shows two predicted simulated concentration profiles across the packed section computed by HYSYS and Aspen. Table 1 provides a comparison between the measured outlet concentrations and the nearest simulation results using the most appropriate thermodynamic package for each case. Although both softwares failed to predict proper outlet concentrations, however, Aspen program provides relatively more realistic results.

Traditional simulation methods such as HYSYS or Aspen software's always use predefined mod-

IJSER

els for their property estimations and they can be modified usually by introducing adequate interactions parameters for a fixed set of components. The selection of proper interaction parameters from mere experimental data can be achieved by resorting to a trial and error procedure which is very time consuming and in most cases it is almost impractical. On the other hand, the empirical models such as RBF or MLP networks are designed to use directly the measured data and provide the required non-linear multivariate interpolation as the output or response.

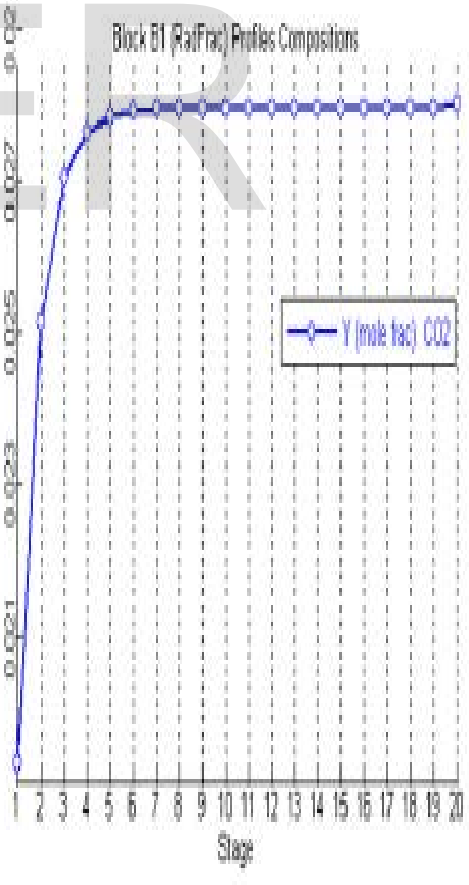
4.1 Neural network simulation results

The entire collection of experimental data presented in Figs. 3 and 4 for CO2 outlet concentrations and percentage of absorbedCO2 were used to train both MLP and RBF networks.

By definition, the RBF regularization network employs the same number of neurons as the data points. For better comparison of both networks performances, the number neurons of MLP networks were kept equal to the number of training exemplars. The regularization network was completely self-sufficient and did not require any initial values for its linear and non-linear

parameters. As it was mentioned in the theoretical section, these networks use all the input data as their centers. Furthermore, our in-house optimization technique was used to select the optimal value of isotropic spreads at each case, separately. The leave one out cross validation (LOOCV or CV in brief) criterion was also employed to provide the optimum level of regularization for a given set of centers and the optimal isotropic spread. In contrast, the MLP networks performances completely depended on the set of initial values selected for their synaptic weights. In this work, various initial synaptic values were tried for each training data set and then the best values were selected based on visual considerations. Evidently, the above procedure for training of MLP networks is much more time demanding than the previously described straightforward method for training of Regularization networks.

Fig. 4 compare the recall and generalization performances of optimally regularized RBF networks with the best



IJSER

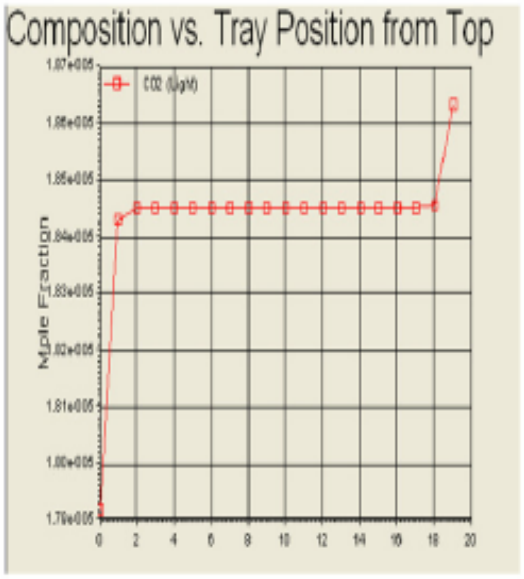


Fig 4. Sample simulation results for packed tower concentration profile (left: Aspen, right: HYSYS).

Table 1

Comparison of HYSYS and Aspen simulation results with measured CO_2outlet concentrations for sample DEA and MDEA solvents.

L (lit/min)	G (lit/min)	Y_i	$Y_{o\ experimental}$	$Y_{o\ HYSYS}$	$\%E_{HYSYS}$	$Y_{o\ Aspen}$	$\%E_{Aspen}$
25 wt % DEA in water							
1	50	0	0	0	*0*	0	*0*
1	50	0.021	0.0013	0.0201	*1446*	0.01397	*975*
1	50	0.0345	0.0074	0.0326	*341*	0.0231	*212*
1	50	0.042	0.0118	0.0402	*241*	0.0281	*138*
1	100	0	0	0	*0*	0	*0*
1	100	0.015	0.0035	0.0144	*311*	0.01222	*249*
1	100	0.029	0.0092	0.0279	*203*	0.0237	*158*
1	100	0.040	0.0144	0.0385	*167*	0.03275	*127*
2	50	0	0	0	*0*	0	*0*
2	50	0.021	0.0009	0.0198	*2100*	0.00719	*699*
2	50	0.0345	0.0061	0.0325	*433*	0.0119	*95*
2	50	0.042	0.0088	0.0396	*350*	0.0146	*66*
2	100	0	0	0	*0*	0	*0*
2	100	0.015	0.0022	0.0143	*550*	0.00993	*351*
2	100	0.029	0.0055	0.0277	*404*	0.0193	*251*
2	100	0.040	0.010	0.0383	*283*	0.0267	*167*
30 wt % MDEA in water							
1	50	0	0	0	*0*	0	*0*
1	50	0.021	0.0056	0.0201	*259*	0.01397	*149*
1	50	0.0345	0.0101	0.033	*227*	0.0231	*129*
1	50	0.042	0.0189	0.0402	*113*	0.0281	*49*
1	100	0	0	0	*0*	0	*0*
1	100	0.015	0.0054	0.0145	*169*	0.01222	*126*
1	100	0.029	0.0112	0.0280	*150*	0.0237	*112*
1	100	0.040	0.0184	0.0386	*110*	0.03275	*78*
2	50	0	0	0	*0*	0	*0*
2	50	0.021	0.0042	0.0198	*371*	0.00719	*71*
2	50	0.0345	0.0079	0.0326	*313*	0.0119	*51*
2	50	0.042	0.0176	0.0397	*126*	0.0146	*–17*
2	100	0	0	0	*0*	0	*0*
2	100	0.015	0.0049	0.0144	*194*	0.00993	*103*
2	100	0.029	0.0104	0.0278	*167*	0.0193	*86*
2	100	0.040	0.0171	0.0383	*124*	0.0267	*56*

5. CONCLUSION

The experimental data collected in this article was used as an influential tool to improve our understandings of the packed absorption processes especially the absorption of CO2 from air by various alkanolamine solutions. The simulation results presented here show that both HYSYS and Aspen software may perform in adequately or prediction of CO2 concentrations profiles across the packed column. This failure emphasizes the complex behavior of the process. As an alternative approach, the performances of two well known classes of artificial neural networks were compared together using the collected experimental data. It was clearly illustrated that the Regularization networks provide more reliable generalization performances compared to MLP networks. The superior performance of RBF networks was due to their solid theoretical foundations and their strong noise filtering capabilities.

6. ACKNOWLEDGEMENT

This work was done by Mr. Sandeep Palve and Mr. Nilesh Kshirsagar Students of Final Chemical Engineering during June 2012 to March 2013, at the Department of Chemical Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai, India Professor S. M. Walke. We are thankful to the authorities at BVCOE for providing all facilities for completing this work.

IJSER

7. REFERENCES

1. Brettschneider, O., Thiele, R., Faber, R., Thielert, H.,Woznya, G., 2004. Separation andPurification Technology 39, 139&159.
2. Datta, A.K., Sen, P.K., 2006. Optimization of membrane unit for removing carbondioxide from natural gas. Journal of Membrane Science 283, 291&300.
3. Fauth, D.J., Frommell, E.A., Hoffman, J.S., Reasbeck, R.P., Pennline, H.W., 2005.Eutectic salt promoted lithium zirconate: novel high temperature sorbent for CO2 capture. Fuel Processing Technology 86, 1503&1521.
4. Golub, G.H., Van Loan, C.G., 1996. Matrix Computations, third ed. Johns HopkinsUniversity Press, Baltimore.
5. Golub, G.H., Heath, M., Wahba, G., 1979. Generalized cross validation as a methodfor choosing a good ridge parameter. Techno Metrics 21 (2), 215.
6. Gray, M.L., Soong, Y., Champagne, K.J., Pennline, H., Baltrus, J.P., Stevens Jr., R.W.,Khatri, R., Chuang, S.S.C., Filburn, T., 2005. Improved immobilized carbon
7. Dioxides capture sorbents. Fuel Processing Technology 86, 1449&1455.
8. Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models, first ed. Chapmanand Hall, London.
9. 9.Haykin, S., 1999. Neural Networks: A Comprehensive Foundation, second ed.Prentice Hall, New Jersey.
10. Huttenhuis, P.J.G., Agrawal, N.J., Hogendoorn, J.A., Versteeg, G.F., 2007. Gas solubilityof H_2S and CO_2 in aqueous solution of N-methyldiethanolamine. Journal of Petroleum Science and Engineering 55, 122&134.

IJSER