Prediction of VLE Data for Ethanol-Water Systems Using Adaptive Neuro Fussy Inference System

A. K. Babayemi[1] and E. E. Eluno[2]
[1,2] Department of Chemical Engineering, Chukwuemeka Odumegwu Ojukwu University,
Anambra State, Nigeria.

ABSTRACT
The prediction of VLE data by conventional thermodynamic methods is tedious and requires determination of "constants" which is arbitrary in many ways hence there is need to adopt Adaptive Neuro Fuzzy Inference System with its associative property and its ability to learn and recognize highly nonlinear and complex relationships. In this work to create a model for prediction of Vapor Liquid Equilibrium data for ethanol water system, the UniSim Process Design Software was used to simulate VLE data for three different ethanol water molar mixtures. An ANFIS model having 5 layers with 9 hidden neurons representing the fuzzy logic rules was thereafter developed to predict the vapor phase fraction of the system. The inputs to the model were the temperature and pressure values simulated on the UNISIM software and the model was trained using input triangular membership function, output linear membership function and the hybrid optimizer to predict the mole fraction and vapor phase fraction of the system taken as outputs. The performance of the ANFIS model gave R-square values of 0.9927, 0.9999 and 0.9999 for the three molar mixtures studied. A comparative plot for the prediction showed a perfect fit for all the three systems for the ANFIS predicted values. It can therefore be concluded that ANFIS model gives a superior predictive capability than conventional models.

Keywords: Prediction, Vapor Liquid Equilibrium, Adaptive, Ethanol-Water.

1.0 INTRODUCTION
A detailed knowledge of the composition of co-existing phases at equilibrium and the effects of temperature and pressure on this composition is required for the development, design and operation of equilibrium stage processes (Kister, 1992). It is difficult and laborious to obtain such data experimentally making such knowledge scarce and unavailable when new systems are under consideration. The ways and means often provided by classical thermodynamics of solutions for treating and testing experimental data are of very little use for process evaluation and design. Predictive methods are therefore valuable for process evaluation and design. Liquid state is highly complex for any general description, being in the in between ' state of 'orderly solids and disorderly gases. Our lack of knowledge of the liquid state precludes the prediction of solution properties from those of the pure components which are, in most cases, either readily available or easily estimated, and forces us to develop empirical methods (Nannoolal et al., 2009).

With the rapid increase in the number and quantity of solvent type compounds, there is concentrated attention on matters relating to vapor-phase fraction in particular. Vapor-phase fraction data are studied primarily with either one or both of the following objectives in mind:
(1) To collect the necessary equilibrium data for design of chemical process equipment.
(2) To provide the essential data needed to formulate models for the prediction of solution properties and test the models.
To date, there has been virtually no first principles-based predictive method that can deal with the problem efficiently with an accuracy surpassing the widely utilized highly successful correlated models such as UNIQUAC Functional Group Activity Coefficients (UNIFAC) (Neiman et al., 2004). In spite of intensive theoretical efforts over the last few years to develop various physical and computational models to predict Vapor Phase Fraction (VPF), it appears that much more remains to be done in order to make the prediction reliable with accuracy comparable or superior to the experimental measurements (Neiman et al., 2004).

Thus, it is necessary to reduce and correlate the limited data to the best possible interpolations and extrapolation. This is the incentive for the application of various correlations to the calculation of phase equilibrium relationships. The Equations of State (EOS) fairly predicts the Vapor Liquid Equilibrium (VLE) data of hydrocarbon systems but is quite handicapped for systems containing polar compounds (Kontogeorgis and Folas, 2010). Further, the EOSs are neither able to describe the critical region satisfactorily for mixtures nor estimate the liquid properties accurately. Activity coefficients are generally used for determining liquid properties and several estimation techniques exist in the literature (Smith & Van Ness, 1995; Fredenslund et al., 1977). However, each has its limitations in its applicability to different systems; for instance, even though the UNIFAC method (Fredenslund et al., 1977) is applicable even to systems containing water, it has its limitations with regard to hydrocarbon mixtures. So, thermodynamics of mixtures are more complicated than for pure compounds and the difficulty in mixture analysis increases with the extent of non-ideality. To speed up flash calculation, attentions have been paid to other numerical tools, such as sparse grids technology and artificial neural networks (ANN). Sparse grids technology is often considered preferable in coupling with flow as the surrogate model created in the offline phase can be used repeatedly (Yuanqing, 2015). However, the generation of the surrogate model is still time- consuming. On the other hand, due to its ability to capture the relationship among large number of variables, especially for non-linear correlations, ANN has attracted considerable attention for performing acceleration. ANNs being purely "numeric" in nature do not require thermodynamic modeling; and hence, are convenient for VLE data prediction (Sharma et al., 1999). A limited database can be used to train a network properly for it to learn the possible pattern of the pressure (P) – temperature (T) – liquid mole fraction (x) – vapor mole fraction (y) surface for a system in a small range of P and T. It has been reported that ANN has been used in the thermodynamic property's calculation successfully, including compressibility factor, vapor pressure and viscosity.

In present work, application of neuro-computing for predicting VLE data will be explored. Conventional thermodynamic techniques for VLE data estimation of mixtures are tedious and have a certain amount of empiricism by way of determining mixture "constants" using arbitrary mixing rules. ANNs, on the other hand, help such predictions and eliminate the need for determining these constants by finding the functional relationship all at once. ANNs also offer the potential to overcome the limitations of existing equations-of- state (EOS) in determining VLE data for highly polar systems. According to Gibbs phase rule, two intensive properties are required to completely describe a binary two-phase system at equilibrium. Thus, two intensive properties can be selected to describe all the properties of the system irrespective of the methodology used. Pressure and Temperature are two convenient intensive properties, since they can be easily measured and controlled. Therefore, the neural equation of state to be developed will use P and T as the inputs and will give mole fraction and molar flow rate of the distillate as the outputs. Ethanol-water system was selected to explore the use of ANNs as an alternative VLE data prediction. Three various molar mixtures were considered for the modeling using ANFIS.

2.0 LITERATURE REVIEW

At thermodynamic equilibrium, the components in a mixture will distribute themselves between the vapor phase and the liquid phase as in Figure 1.

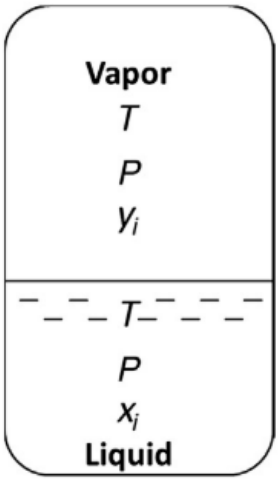

Figure 1: Vapor Liquid Equilibrium.

The distribution of component i between the vapor phase, with composition y_i, and the liquid phase with composition x_i, is normally represented by the distribution coefficient:

$$K_i = \frac{y_i}{x_i} \qquad (1)$$

The distribution coefficient, or K-value, is generally a function of temperature and pressure only and not of compositions. A component with a high K-value will be quite volatile, as the vapor phase will have a high concentration of that component. Similarly, a low K-value means the component is not particularly volatile. For successful separation of a mixture, the main interest is the difference between the volatilities of the components, as a large difference will mean an easy separation.

In order to find the relationship between input and output data driven from accelerated experimentations, a powerful method than traditional modeling is necessary. Artificial Neural Network (ANN) can be cited as 'algorithmic equivalent' of the human learning process and information processing scheme at a modest scale. They are pattern recognition architecture which can identify patterns between complex sets of input and output data (Hagan, M.T., H.B. Demuth, 1996; Biglin, M., 2004). These patterns are then used to predict outcomes for fresh inputs. They do not require the specification of correlations which govern process, but are trained on real life data. The chief advantage of ANNs lies in the fact that ANN uses a generic model which covers a wide class of problems. It does not require a fundamental understanding of the process or phenomena being studied and can handle complex and nonlinear models. The ANNs are able to deal with non-linear problems, and once trained can perform prediction and generalization at high speed (Sozen, A., E. Arcakilioglu, 2004). They have been used to solve complex problems that are difficult for conventional approaches, such as control, optimization, pattern recognition, classification and properties that desired that the difference between the predicted and observed (actual) outputs be as small as possible (Richon, D., S. Laugier, 2003).

Thus, they are gaining a rapid interest within engineering, medical, financial and various other fields. They have thus made strong advances in area of continuous speech recognition, classification of noisy data, market forecasting, process modelling, fault detection and control. Mehmet Bilgin (2004), employed a neural network model to calculate the isobaric vapour-liquid equilibrium of binary systems composed of different chemical structures, which do not show azeotropic behavior. Results generated by the ANN model were compared to those generated by the UNIFAC and the Margules model. In all cases, the deviations between the experimental activity co-efficients and those calculated by the neural network were less than those obtained by those obtained

by the Margules and the UNIFAC models. Moghadassi et al., (2010), developed a model for predicting the VLE data for binary systems containing propane. Four binary refrigerant systems containing propane were considered. Results generated by the ANN model were compared with those generated by Margules and Vaan Laar models. The ANN model showed superiority over the other thermodynamic models. Pandharipande et.al (2012) developed a model for the evaluation of VLE data for ten binary systems, results obtained indicated minimum error is obtained in the case of ANN models. Pandharipande et.al, (2012), modeled combined VLE of four quaternary mixtures using artificial neural network. It was observed that in the ANN model, the error difference between the predefined value and output calculated is minimized. Nasri et al., (2012) developed an ANN model to predict the VLE of a carbon dioxide methanol system at high pressure. Predicted values using ANN are satisfactory.

A general overview of the literature shows that the experimental data or a set data points (experimental) available in literature are fed to the neural network for pattern recognition. Pattern recognition or the relation between the input and output is generated by the ANN. Thus, in other words, the ANN generates a mathematical model. Once this model is generated, it is tested for its accuracy with a set of new data points which have not been fed earlier to the ANN for pattern recognition. By entering a new set of data points for testing, the efficiency of prediction of the developed ANN model can be known.

3.0 MATERIALS AND METHOD

The materials utilized for this study were mainly software projected with a personal which include:

- UNISIM: A software widely used by processing engineers for the design of chemical engineering process routes, schedules of the equipment design and processing of raw materials into finished goods (steady state and dynamic process).
- MATLAB: mainly used for mathematical computation and modeling. It is vast in various areas of model generation, optimization and artificial intelligence modeling.
- MICROSOFT WORD: used to create professional looking documents.

The vapor-phase fraction, the molar flow rate and the mole fraction of the volatile component were obtained using UNISIM for the three various molar mixtures of ethanol and water. The dependent variables (vapor phase fraction, mole fraction and molar flow of ethanol for the product stream of the flash column) were determined at varying temperature and pressure. An artificial intelligence model was developed to predict the dependent variables with temperature and pressure data as the independent variables. MATLAB was employed in performing the empirical and intelligent modeling process. The Adaptive Neuro Fuzzy Inference System (ANFIS) (being the most accurate model for use in terms of prediction accuracy as proposed by previous studies in various areas) was used in this research as the artificial intelligence model. On completing the prediction by the two models, prediction performance parameters and error statistics were used to determine the model with more prediction accuracy and less error values.

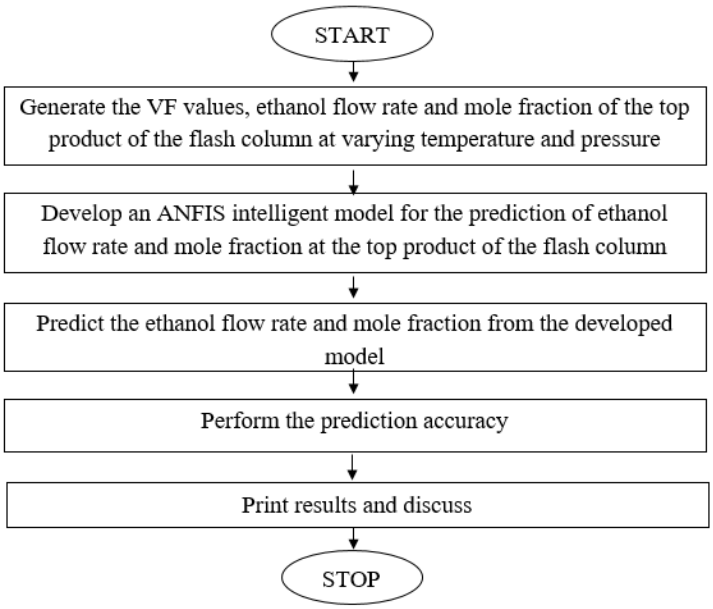

Figure 2: Flow chart for the research procedure.

Data Acquisition

The parameters necessary to obtain the vapor phase fraction (VPF), ethanol flow rate and mole fraction from UNISIM software are as outlined:

- ✓ Molar mixture of ethanol and water of 40:60, 60:40 and 45:55 respectively.
- ✓ Molar flow rate of the stream being 100kgmole/hr.
- ✓ Varying temperature and pressure values.

The temperature and pressure values for the simulation using the UNISIM software are shown in Table 1 to Table 3.

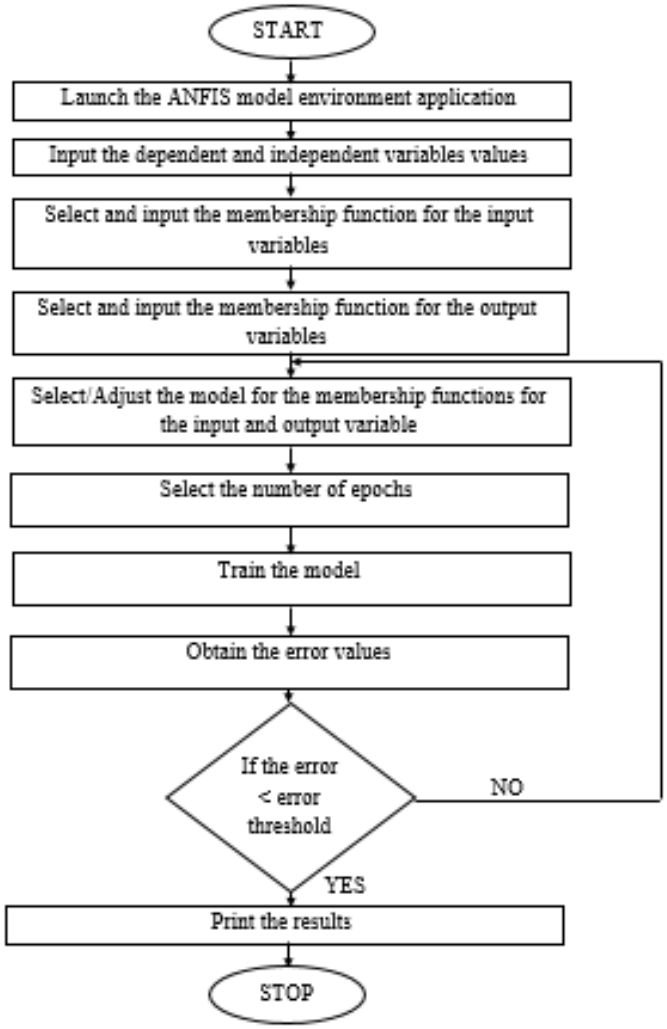

Figure 3: Flow chart for ANFIS modeling.

4.0 RESULTS AND DISCUSSION

In this work to predict and optimize the flash calculation, an ANFIS was used to model input and output process of the NPT flash column for an equi-molar ethanol/water system using the linear membership function and a hybrid optimization algorithm. Essentially, as the Gibbs phase rule stipulates, two intensive properties are required to completely describe a binary two-phase system at equilibrium conditions. Temperature and pressure are two such thermodynamic intensive properties conventionally selected as inputs to the model, because of the relative ease with which they can be measured. Molar flow rate and mole fraction values of the distillate for a total of 11 temperature/pressure data points for each molar mixture studied were obtained from the UNISIM software flash simulation for training the ANFIS model. The aim function which was used to select the ANFIS as the best model is the root squared error.

Table 1: Simulation Result for Temperature & Pressure Effects on 40:60 Molar Mixture.

Temperature T (°C)	Pressure P (KPa)	Vapor Phase Fraction VPF	Ethanol flow rate at the top Y_{flow} (kgmole/hr)	Ethanol mole fraction at the top $Y_{f\ rac.}$ (%)
60.6	53.33	0.0000	0.0000	0.0000
69.0	62.00	0.2489	14.8817	59.79
70.0	63.00	0.4251	24.4688	57.56
71.0	64.00	0.5349	29.7029	55.53
75.0	80.00	0.1844	11.1193	60.30
77.5	87.00	0.3235	18.9959	58.72
80.0	95.00	0.3919	22.6283	57.74
81.0	100.00	0.3118	18.3245	58.77
83.7	105.00	0.5720	31.0310	54.25
88.0	120.00	0.6510	33.8455	51.99
89.0	125.00	0.6419	33.5264	52.23

Table 2: Simulation Result for Temperature Pressure Effects on the 60:40 Molar Mixture.

Temperature (°C)	Pressure (KPa)	Vapor Phase Fraction	Ethanol flow rate at the top (kgmole/hr)	Ethanol mole fraction at the top (%)
50.0	45.00	0.0000	0.0000	0.0000
62.7	50.70	0.2337	16.0879	68.84
65.8	58.00	0.2632	18.0660	68.64
67.6	61.00	0.7444	47.9989	64.48
69.0	62.00	1.0000	60.0000	60.00
71.0	64.00	1.0000	60.0000	60.00
72.5	74.00	0.8560	53.8338	62.89
74.6	80.00	0.9097	56.3650	61.96
75.7	83.50	0.9203	56.8377	61.76
77.5	87.00	1.0000	60.0000	60.00
80.0	95.00	1.0000	60.0000	60.00

Table 3: Simulation Result for Temperature Pressure Effects on the 45:55 Molar Mixture.

Temperature (°C)	Pressure (KPa)	Vapor Phase Fraction	Ethanol flow rate at the top (kgmole/hr)	Ethanol mole fraction at the top (%)
67.6	61.0	0.0000	0.0000	0.00
69.0	62.00	0.4386	26.2238	59.79
70.0	63.00	0.5889	33.8971	57.56
71.0	64.00	0.6847	38.0214	55.53
72.5	74.00	0.1097	6.8990	62.89
74.6	80.00	0.2188	13.5568	61.96
77.5	87.00	0.5042	29.6066	58.72
80.0	95.00	0.5633	32.5249	57.74
83.7	105.00	0.7222	39.1794	54.25
88.0	120.00	0.7965	41.4100	51.99
91.0	128.00	0.8979	43.5930	48.55

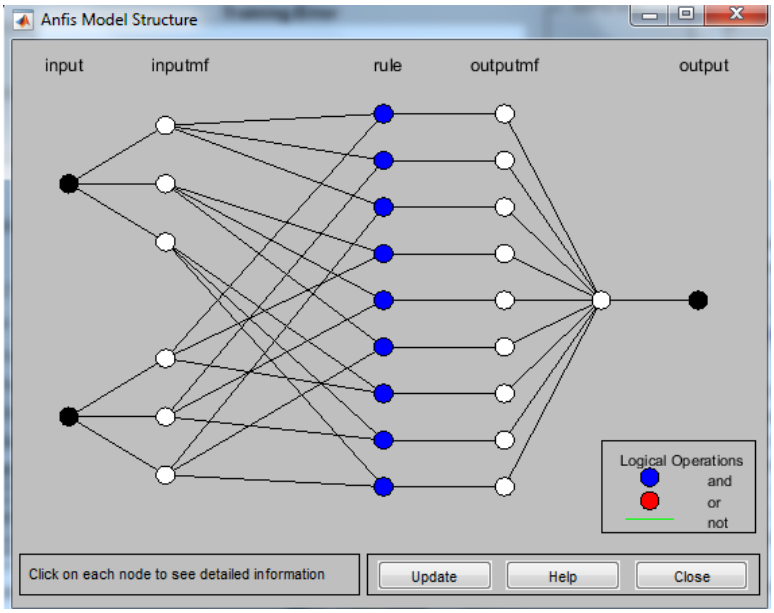

Figure 4: ANFIS architecture for the model predictions.

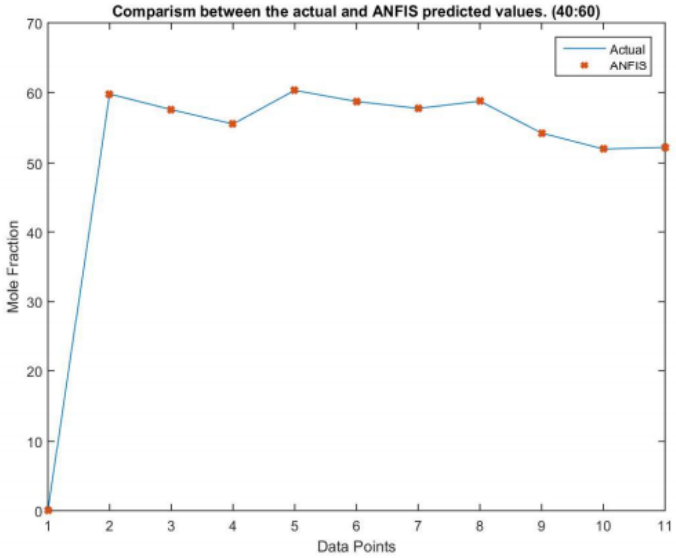

Figure 5: Comparative plot for the ANFIS model prediction of the 40:60 molar mixture.

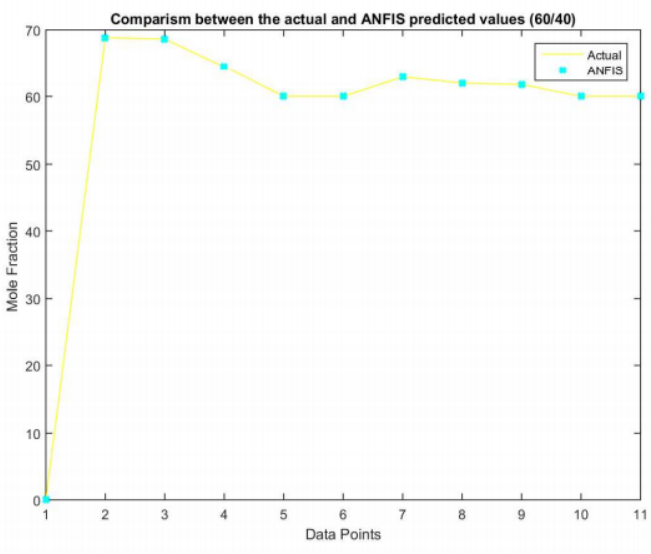

Figure 6: Comparative plot for the ANFIS model prediction of the 60:40 molar mixture.

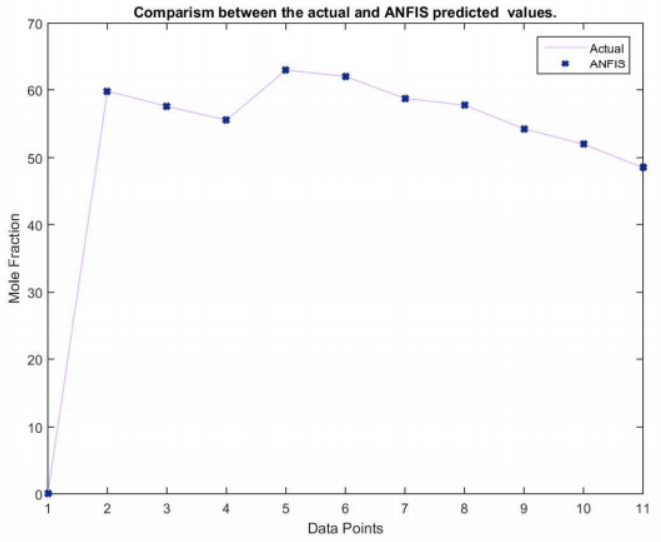

Figure 7: Comparative Plot for the ANFIS Model Prediction of the 45:55 Molar Mixture

Table 4: Prediction Performance Parameter for the Model Predictions

Molar Ratio of Mixture	R^2
40:60	0.9927
60:40	0.9999
45:55	0.9999

Data Selection

The NPT flash simulation was done with 11 data points of variable temperature and pressure points as presented in Table 1 to Table 3 for the three various molar compositions. The temperature range within the boiling points of the components of the mixture (78^{o}C and 100^{o}C for ethanol and water respectively) and the pressure range that gave tangible mole fractions of ethanol at the top were considered. Notwithstanding the stochastic nature of the data set, it was expected of the model to give a robust performance after training as ANFIS are proven to model even stochastic data with great accuracy.

The ANFIS Model

The ANFIS modeling was a supervised learning algorithm as the algorithm was provided with not only the inputs, but also the targets (desired outputs).An input triangular membership function and a linear membership function for the output were chosen in generating the fuzzy system. As shown in Figure 4, the model took the two variables (temperature and pressure) as inputs and the value of mole fraction for the distillate as the output. The key parameters of the model are the weights of each layer, which determines the outputs from the model given the input. At first, those weights were initialized randomly, which meant that the model gave insignificant output values with the given inputs. To make the model useful for this problem, those weight parameters were trained with the fuzzy inference system to fit our problem. In doing so, an iterative optimization process was run on the model's parameters to make the loss (error) as small as possible. The difference between the model's output and the expected value is referred to as loss. Here, for this regression problem, R-squared was used as the loss function. As the training progressed, the loss of the model on the validation dataset decreased gradually,

which implies that the difference between the model's output and the expected value became smaller, and the model became increasingly useful.

Table 1 to Table 3 show the intelligence model prediction results for the mole fractions of the distillate for the various temperature/pressure data sets. The predictions for the data sets followed similar trajectory of the actual data gotten from UNISIM simulation for the mole fraction values for the three systems studied. To demonstrate the relative accuracy of the ANFIS model with UNISIM data (seen as real data), Figure 5 to Figure 7 show the comparative plot for the actual and predicted values of the mole fraction of the distillate for the various molar mixtures studied. The intelligence model gave very robust predictions of the mole fractions as seen from the figures. Also, the very high prediction coefficient of determination (R-square) of 0.99 as seen in Table 4 validates the robustness of the ANFIS model in predicting the thermodynamic properties of interest. The developed ANFIS model helped such predictions and eliminated the need for determining the various thermodynamic constants required by conventional models by finding the functional relationship all at once. As a result, the machine learning method was used to achieve a robust model to represent the properties.

CONCLUSION

An ANFIS model has been developed for the prediction of VLE data in a non-linear binary system (ethanol/water) and the model predicts the composition and flow rate of ethanol in vapor phase. A set of 11 experimental data points for three various molar mixtures were used for the network training. The best architecture is the triangular input and linear output membership functions obtained through trial and error. The performance of the proposed ANFIS model was also tested through its application to a checking data set after the training. The results of this estimation indicate that the developed ANFIS model will be able to predict VLE data with very high accuracy. It will also be useful in 'on the spot' monitoring and control in industrial VLE applications. An even spread of the data seemed to be essential for better predictions. Thus, in principle ANFIS hold promise as a strategy for solving the tedious VLE data generation problem.

REFERENCES.

Fredenslun A., Gmehling J., and Rasmussen P., (1997). *Vapor Liquid Equilibrium Using UNIFAC*, Elsevier, Amsterdam.

Hagan M.T., Menhaj M. (1994). *Training feed forward networks with the Marquardt algorithm"*, IEEE Trans. Neural Networks, 5(6), 989--993.

Kister, Henry Z. (1992). *Distillation Design* (1st ed).McGrawHill. ISBN 978-0-07-034909-4.

Kontogeorgis G.M., Folas G.K. (2010).*Thermodynamic Models for Industrial Applications. From Classical and Advanced Mixing Rules to Associating Theories.* Ed. Wiley, Chichester, West Sussex, United Kingdom.

Laugier S., Richon D.,(2003)*Use of artificial neural networks for calculating derived thermodynamic quantities from volumetric property data*, Fluid Phase Equilibria, Volume 210, Issue 2, Pages 247-255, ISSN 0378-3812, https://doi.org/10.1016/S0378-3812(03)00172-9.

Mehmet B., Hasdemir H. I., Oztas O., (2004) Journal of scienctific and industrial Research, 63, 336.

Nannoolal, Yash & Rarey, Jürgen & Ramjugernath, Deresh. (2009). *Estimation of pure component properties. Part 4: Estimation of the saturated liquid viscosity of non-electrolyte organic compounds via group contributions and group interactions.* Fluid Phase Equilibria. 281. 97-119. 10.1016/j.fluid.2009.02.016.

Nasri, Fariborz & Daraei, Hiua & Hatami, Tahmasb & Maleki, Afshin. (2012). *Phase Equilibrium of Binary System Carbon Dioxide-Methanol at High Pressure Using Artificial Neural Network.* Journal of The Chemical Society Of Pakistan. 34. 1070-1078.

Neiman, M., Cheng, H., Parekh, V., Peterson, B., & Klier, K. (2004). *A critical assessment on two predictive models of binary vapor–liquid equilibrium. Phys. Chem. Chem. Phys., 006(13), 3474–3483.* doi:10.1039/b400322e

Nikkholgh, M.R., Moghadassi,A.R., Parvizian,F., and Hosseini,S.M., (2010) Estimation of vapour–liquid equilibrium data for binary refrigerant systems containing 1,1,1,2,3,3,3heptafluoropropane(r227ea) by using artificial neural networks. The Canadian Journal of Chemical Engineering, 88(2):200–207.

Pandharipande, Shekhar and Shah, Anish M, (2012). *Modeling combined VLE of four quaternary mixtures using artificial neural network.* International Journal of Advances in Engineering, Science and Technology (IJAEST), 2,2, 169-177
Sharma, R., Singhal, D., Ghosh, R., Dwivedi, A., (1999) Computers and Chem. Engng. 23 385.

Smith , J. M., & Van Ness, H. C., (1995). *Thermodynamics for Chemical Engineers* (5th ed.). New York: McGraw Hill.

Sozen, A., E. Arcakilioglu, M. Ozalp, 2004. Investigation of thermodynamic properties of refrigerant /absorbent couples using artificial neural networks, Chemical engineering and processing, 43: 1253-1264.

Yuanqing Wu, Christoph Kowitz, Shuyu Sun, and Amgad Salama. (2015) *Speeding up the flash calculations in two-phase compositional flow simulations–the application of sparse grids.* Journal of Computational Physics, 285:88–99.

MATLAB and Statistics Toolbox Release 2015a, The MathWorks, Inc., Natick, Massachusetts, United States.