ORIGINAL ARTICLE

Correlation of ternary liquid–liquid equilibrium data using neural network-based activity coefficient model

Atilla Özmen

Abstract Liquid–liquid equilibrium (LLE) data are important in chemical industry for the design of separation equipments, and it is troublesome to determine experimentally. In this paper, a new method for correlation of ternary LLE data is presented. The method is implemented by using a combined structure that uses genetic algorithm (GA)–trained neural network (NN). NN coefficients that satisfy the criterion of equilibrium were obtained by using GA. At the training phase, experimental concentration data and corresponding activity coefficients were used as input and output, respectively. At the test phase, trained NN was used to correlate the whole experimental data by giving only one initial value. Calculated results were compared with the experimental data, and very low root-mean-square deviation error values are obtained between experimental and calculated data. By using this model tie-line and solubility curve data of LLE can be obtained with only a few experimental data.

Keywords LLE · Neural network · Genetic algorithm · Activity coefficients

1 Introduction

In the chemical process industries, fluid mixtures are often separated into their components by diffusional operations such as distillation, absorption, and extraction. Design of such separation equipments requires quantitative estimates

A. Özmen (✉)
Department of Electrical-Electronics Engineering,
Faculty of Engineering and Natural Sciences,
Kadir Has University, Cibali, 34083 Istanbul, Turkey
e-mail: aozmen@khas.edu.tr

of the partial equilibrium properties of fluid mixtures [1]. The partial equation properties of liquid mixtures are mainly presented with LLE data. For example, an organic phase is required in order to separate a chemical component from an initial aqueous solution via extraction methods. In this case, at the end of the process, the component to be separated has got LLE data in both organic and aqueous phases. These LLE data can be either determined experimentally or predicted/estimated by well-known thermodynamics models. Experimentally determination of LLE data is time and energy consumed. On the other hand, the conventional model equations (NRTL, Margules, UNIFAC, UNIQUAC) do not give always accurate results for each liquid mixture, where infinite variation of chemical components is available.

In the literature, LLE data are generally estimated using thermodynamic models based on the well-known fundamental phase equilibrium criterion of equality of chemical potential in both phases. These models are called generally as "activity coefficient models," and many empirical and semitheoretical equations exist for estimating activity coefficients of binary mixtures containing polar and/or nonpolar species [2].

Different methods have been suggested in the literature for phase equilibrium calculations. In the literature, NN and GA usually have been used for estimating vapor–liquid equilibria (VLE) [3–5]. GA has been also used to obtain the interaction parameters of known methods such as NRTL and UNIQUAC [6]. In these methods, activity coefficient models are used to obtain LLE. The aim of this work is to develop a new model which uses only a few experimental data for estimating the whole LLE data for a wide range of mixture variations. Experimental equilibrium data for these systems were taken from literature [7]. The method estimates LLE data by using a combined structure that consists

of NN and GA. NN structure that is used in this work is multilayer feed-forward network with one hidden layer. In the literature, H. Ghanadzadeh's paper uses a different NN structure that is called GMDH type NN which is based on Kolmogorov–Gabor polynomial [8].

2 Phase equilibria problem

2.1 Activity and activity coefficient

Activity is a measure of the "effective concentration" of a species in a mixture. Activity and activity coefficient and their estimation are important concepts for equilibrium calculations. The activity of component i at some temperature, pressure, and composition is defined as the ratio of the fugacity of i at these conditions to the fugacity of i in the standard state, that is a state at the same temperature as that of the mixture and at some specified condition of pressure and composition [9].

$$a_i(T,P,x) \equiv \frac{f_i(T,P,x)}{f_i(T,P^0,x^0)} \tag{1}$$

where P^0 and x^0 are, respectively, an arbitrary but specified pressure and composition. The fugacity may be looked upon as a sort of corrected pressure that will describe the behavior of an actual gas in the manner of an ideal gas [10].

The activity coefficient γ_i is the correction factor which measures the departure of a solution from ideal behavior for a given standard state. It is the ratio of the activity of i to some convenient measure of the concentration of i, usually the mole fraction [11].

$$\gamma_i \equiv \frac{a_i}{x_i} \tag{2}$$

In the liquid mixture, all activity coefficients are directly related to the molar excess Gibbs energy G^E which is defined by

$$G^E = RT \sum_i x_i ln\gamma_i \tag{3}$$

where R is gas constant (8.314 J mol^{-1}K^{-1}) and T is absolute temperature in Kelvin. A mathematical model, preferably based on molecular considerations, provides a convenient method for expressing G^E as a function of x. From this function an individual activity coefficient γ_i for component i can be calculated from G^E [12].

2.2 Activity coefficient models

In a liquid system composed of two phases, the Gibbs energy of mixing of the individual phases, ΔG, may be expressed as follows:

$$\begin{aligned} &(n^I + n^{II})\Delta G = n^I \Delta G^I + n^{II} \Delta G^{II} \\ &(constraint : n_i^I + n_i^{II} = n_i, i = 1, 2, 3) \end{aligned} \tag{4}$$

where n_i is the total number of moles of component i and ΔG^I and ΔG^{II} are the Gibbs energies of mixing corresponding to n^I moles of phase I and n^{II} moles of phase II.

The molar Gibbs energy of mixing for either phase I or II is the sum of the ideal and the excess molar Gibbs energies of mixing:

$$\Delta G^I = G_{id}^{(I)} + G^{E(I)} \tag{5}$$

$$= RT \sum x_i^I ln x_i^I + RT \sum x_i^I ln\gamma_i^I \tag{6}$$

$$= RT \sum x_i^I ln a_i^I \tag{7}$$

where γ_i is the activity coefficients and a_i are activities.

The necessary and sufficient criterion of equilibrium is that ΔG for the system is minimum. Since ΔG is minimum, a differential change of composition occurring at equilibrium at fixed pressure and temperature will not produce any change in ΔG and hence: $d(\Delta G)_{P,T} = 0$.

This criterion is a necessary, but not sufficient condition of equilibrium. It does not help us in distinguishing between a maximum, an inflection point, and a minimum. The aforementioned necessary, but not sufficient condition, may be stated in an alternative way: The activity a_i for each component must be the same in the two phases:

$$\begin{aligned} &a_i^I = a_i^{II} \quad i = 1, 2, 3 \\ &(constraint : \sum x_i^I = \sum x_i^{II} = 1 \quad i = 1, 2, 3) \end{aligned} \tag{8}$$

This criterion may be easy to use in practice, but it suffers from not being sufficient. The maxima and the inflection points may be avoided by a check for convexity of the predicted $\Delta G -$ curve at the concentrations which are found from the isoactivity criterion

$$x_i^I \gamma_i^I = x_i^{II} \gamma_i^{II} \quad i = 1, 2, 3 \tag{9}$$

where γ_i is the activity coefficient of component i [13].

3 Neural network

NNs were inspired by the power, flexibility, and robustness of the biological brain. They were computational analogs of the basic biological components of a brain (i.e., neurons, synapses, and dendrites). NNs consist of many simple mathematical elements that work together in parallel and in series. A NN model can be seen in Fig. 1. Each neuron has many inputs and only one output, and this output is the input of the other neurons.

As shown in Fig. 2, a neuron model consists of a summing junction and an activation function. Here, $x_1, x_2, x_3, \ldots x_n$ are inputs; $w_1, w_2, w_3, \ldots w_n$ are weight
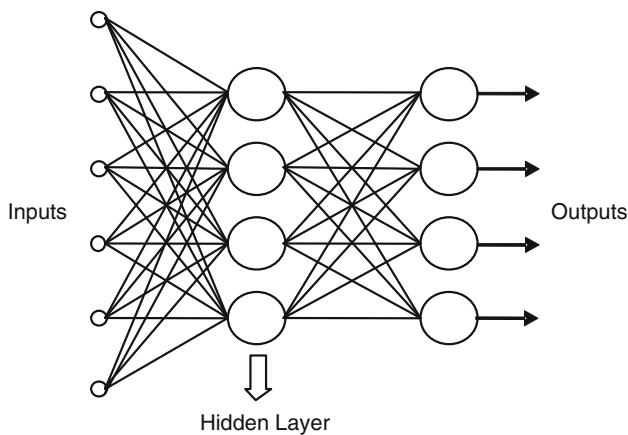
Fig. 1 Neural network model with input, hidden, and output layers

coefficients; b is the bias and y is the output. In this model, the output equation can be given as the following:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{10}$$

where $f(x)$ is activation function [14].

3.1 Training process

A training process can be viewed as the problem of determining network architecture and weight coefficients so that neural network can perform a special task. Learning rule is an equation set by which all or some of the weight coefficients change so as to modify the response of each neuron in time. By this way, NN can adapt itself to get the desired response.

NNs are learnt by example data instead of programming. The network usually must learn the weight coefficients
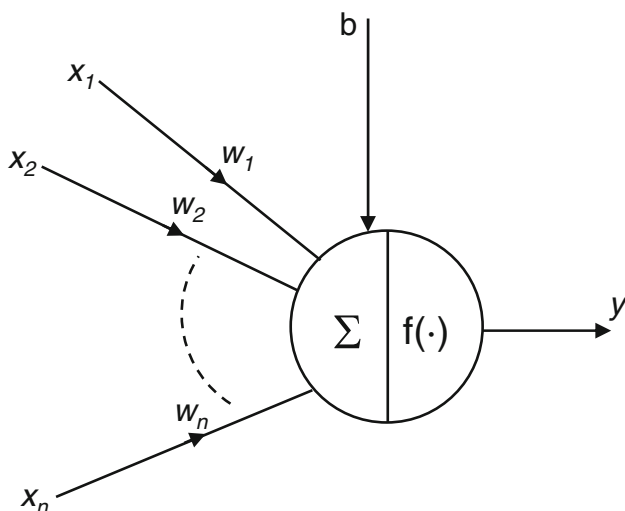

Fig. 2 Neuron model with input and output

from available training set. Learning process can be divided into two groups: supervised and unsupervised learning. In supervised learning, both the input and the response are given to the system. For each input, obtained response and desired response are compared. To get the minimum difference, weight coefficients are changed. After an acceptable error is obtained, learning process is stopped and then these weight coefficients can be used with the data that are not used in learning process. NN begins in a random state and learns using repeated processing of a data training set, which is a set of inputs with target outputs. Learning process occurs because the error between NN output and the target output is calculated and used to adjust the weighted synapses of the NN. This continues until errors are small enough or no more weight changes are occurring. Thus, NN is trained and the weights are fixed. The trained NN can be used for new inputs to perform estimation or classification of tasks [15].

4 Genetic algorithm

Genetic algorithm is a heuristic search algorithm that is inspired by the biological evolution process and used to find the solution of the optimization problems. Algorithm is started with a set of possible solutions. This set of solutions is called as population and represented by chromosomes [16].

One common application of genetic algorithm is function optimization, where the goal is to find a set of parameter values that maximize a multiparameter function (fitness function).

4.1 Genetic algorithm operators

The basic form of genetic algorithm consists of three types of operators: selection, crossover, and mutation. Selection operator selects chromosomes in the population for reproduction. Crossover decomposes two distinct solutions and then randomly mixes their parts to form new solutions. Mutation randomly alters some of gene values in a chromosome from its initial state. This operation results in new gene values, and better solution values can be obtained from this new gene values.

4.2 Basic genetic algorithm

A simple genetic algorithm works as follows:

- Start with a randomly picked population (candidate solutions to a problem).
- Calculate the fitness values of each chromosome in the population.

- Create new population by repeating the following steps:
  - Using fitness probabilities select a pair of parent chromosomes from the current population.
  - Cross over the pair at a randomly chosen point to form two offspring.
  - Mutate the two offspring and place the obtained chromosomes in the new population.
- Replace the current population with the new population.
- Repeat process by starting from fitness value calculation step [17].

5 Proposed NN model

The proposed NN model was used to estimate the activity coefficients of ternary LLE data. Experimental data were used for the input of the network, and corresponding activity coefficients have been selected for the output as shown in Fig. 3. The proposed NN structure is a multilayer feed-forward network with one hidden layer. In NN, hidden layers are layers that connect input to output via a set of neurons. In this work for each of experimental sets, five neurons were used in the hidden layers. Since activity coefficient model has a nonlinear structure and to introduce this nonlinearity into NN, hyperbolic tangent sigmoid function was used in the hidden layer as activation function that is given in Eq. (11). In the output, no activation function was used.

$$f(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \qquad (11)$$

In this figure x_1, x_2, and x_3 represent the experimental data sets of solvent or water phase and γ_i shows corresponding activity coefficients. Weight coefficients of NN
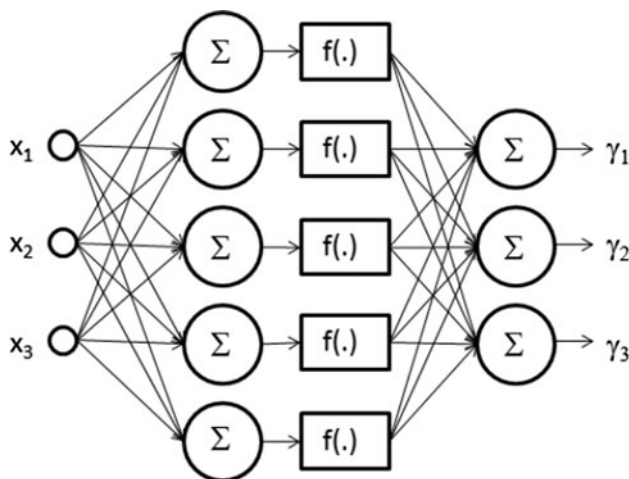

Fig. 3 Proposed NN model

were determined numerically by using a hybrid method that uses genetic algorithm and a search algorithm by minimizing the following objective function.

$$F_a = \sum_{j=1}^{N} \sum_{i=1}^{3} \left(x_{ij}^I \gamma_{ij}^I - x_{ij}^{II} \gamma_{ij}^{II} \right)^2 \Big/ \left(x_{ij}^I \gamma_{ij}^I + x_{ij}^{II} \gamma_{ij}^{II} \right)^2 \qquad (12)$$

where x_{ij}^I and x_{ij}^{II} stand for the experimental mole fraction of component i in water-rich and solvent-rich phases, respectively, along tie-line, j, γ_{ij}^I and γ_{ij}^{II} are the corresponding activity coefficients calculated by NN model, and N is the total number of tie-lines [13].

For minimization, MATLAB optimization toolbox was used. First, genetic algorithm was used to obtain initial values, and then search function was used to obtain final values. By using this hybrid search method for each data set, objective function minimization process was repeated until no further minimization is possible. After training process, any single mole fraction of water, in water-rich phase including experimental values, was used for testing.

Test process was carried out as follows: Firstly, any single mole fraction value of water that randomly picked from water-rich phase (x_1^I) was taken as known value. Then, the mole fraction of acid in water phase (x_2^I) and mole fractions of water (x_1^{II}) and acid (x_2^{II}) in solvent-rich phase were selected as free parameters. Then by using these mole fractions, activity coefficients of water and solvent phase mole fractions were obtained by the trained NN structure. Finally, the following objective function was minimized to obtain the tie-lines. If k shows any tie-line that corresponds to the known water data in the water phase, the objective function F_b was minimized under the given constraints.

Table 1 Model tie-line data for water (1) + acetic acid (2) + dimethyl maleate (3) at $T = 298.2$ K

Water-rich phase			Solvent-rich phase		
x_1^I	x_2^I	x_3^I	x_1^{II}	x_2^{II}	x_3^{II}
0.9800	0.0091	0.0109	0.1420	0.0392	0.8188
0.9700	0.0170	0.0130	0.1938	0.0622	0.7441
0.9600	0.0253	0.0147	0.2529	0.0827	0.6644
0.9500	0.0332	0.0168	0.3137	0.0995	0.5868
0.9400	0.0403	0.0197	0.3730	0.1117	0.5154
0.9300	0.0468	0.0232	0.4315	0.1197	0.4488
0.9200	0.0529	0.0271	0.4910	0.1240	0.3850
0.9100	0.0589	0.0311	0.5528	0.1251	0.3221
0.9000	0.0651	0.0349	0.6160	0.1235	0.2605
0.8900	0.0714	0.0386	0.6765	0.1198	0.2037
0.8800	0.0776	0.0424	0.7278	0.1150	0.1573
0.8700	0.0832	0.0468	0.7661	0.1101	0.1238
0.8600	0.0878	0.0522	0.7930	0.1058	0.1011
0.8500	0.0914	0.0586	0.8100	0.1023	0.0877

$$F_b = \sum_{i=1}^{3} \left(x_{ik}^I \gamma_{ik}^I - x_{ik}^{II} \gamma_{ik}^{II}\right)^2 / \left(x_{ik}^I \gamma_{ik}^I + x_{ik}^{II} \gamma_{ik}^{II}\right)^2$$

$$\text{constraints}: x_{1k}^I + x_{2k}^I + x_{3k}^I = 1 \quad x_{1k}^{II} + x_{2k}^{II} + x_{3k}^{II} = 1 \quad (13)$$

$$x_{1k}^I \geq 0 \quad i = 1,2,3$$

$$x_{1k}^{II} \geq 0 \quad i = 1,2,3$$

Minimization was obtained in two steps. In the first step genetic algorithm was used to obtain initial condition of the

second step. In the second step a MATLAB function, named fminsearch, was used to get the solution that gives minimum RMSD errors. fminsearch uses the simplex search method of Lagarias [18]. This is a direct search method that does not use numerical or analytic gradients. This algorithm can be applied to discrete optimization problems, where derivate-based optimization methods cannot be used.

6 Results and discussion

The benchmarking system involves the liquid–liquid equilibrium data correlations for the ternary systems of water (1) + carboxylic acid (2) + dimethyl maleate (3). Formic, acetic, propionic, and butyric acids were used as carboxylic acid. The calculated tie-line data were obtained by increasing the water mole fraction in steps of 0.01 and shown in Tables 1, 2, 3, 4. The experimental and correlated data of obtained water (1) + carboxylic acid (2) + dimethyl maleate (3) systems have been shown in Table 5 in which x_i^I and x_i^{II} refer to mole fraction of the ith component in the aqueous and solvent phases, respectively. The RMSD values for the studied systems were also listed in

Table 2 Model tie-line data for water (1) + butyric acid (2) + dimethyl maleate (3) at $T = 298.2\ K$

Water-rich phase			Solvent-rich phase		
x_1^I	x_2^I	x_3^I	x_1^{II}	x_2^{II}	x_3^{II}
0.9800	0.0112	0.0088	0.1199	0.0221	0.8580
0.9700	0.0214	0.0086	0.1587	0.0354	0.8059
0.9600	0.0313	0.0087	0.1893	0.0456	0.7652
0.9500	0.0409	0.0091	0.2137	0.0540	0.7323
0.9400	0.0504	0.0096	0.2338	0.0612	0.7050
0.9300	0.0597	0.0103	0.2508	0.0676	0.6816
0.9200	0.0688	0.0112	0.2659	0.0732	0.6609
0.9100	0.0776	0.0124	0.2795	0.0784	0.6420
0.9000	0.0862	0.0138	0.2924	0.0833	0.6243
0.8900	0.0945	0.0155	0.3045	0.0880	0.6075
0.8800	0.1025	0.0175	0.3163	0.0925	0.5912
0.8700	0.1104	0.0196	0.3278	0.0971	0.5751
0.8600	0.1182	0.0218	0.3395	0.1019	0.5586
0.8500	0.1258	0.0242	0.3517	0.1071	0.5412
0.8400	0.1332	0.0268	0.3644	0.1125	0.5231
0.8300	0.1401	0.0299	0.3772	0.1179	0.5049
0.8200	0.1465	0.0335	0.3897	0.1232	0.4871
0.8100	0.1524	0.0376	0.4018	0.1282	0.4699
0.8000	0.1580	0.0420	0.4136	0.1333	0.4532
0.7900	0.1634	0.0466	0.4250	0.1382	0.4368
0.7800	0.1684	0.0516	0.4359	0.1429	0.4212
0.7700	0.1731	0.0569	0.4464	0.1474	0.4062
0.7600	0.1774	0.0626	0.4564	0.1518	0.3918
0.7500	0.1813	0.0687	0.4659	0.1559	0.3781
0.7400	0.1849	0.0751	0.4751	0.1599	0.3650
0.7300	0.1882	0.0818	0.4841	0.1637	0.3523
0.7200	0.1911	0.0889	0.4928	0.1673	0.3399
0.7100	0.1937	0.0963	0.5013	0.1709	0.3278
0.7000	0.1959	0.1041	0.5097	0.1742	0.3161
0.6900	0.1979	0.1121	0.5181	0.1775	0.3044
0.6800	0.1996	0.1204	0.5265	0.1807	0.2928
0.6700	0.2009	0.1291	0.5350	0.1838	0.2813
0.6600	0.2020	0.1380	0.5436	0.1868	0.2696
0.6500	0.2027	0.1473	0.5522	0.1896	0.2582
0.6400	0.2031	0.1569	0.5611	0.1922	0.2466
0.6300	0.2031	0.1669	0.5703	0.1947	0.2350

Table 3 Model tie-line data for water (1) + formic acid (2) + dimethyl maleate (3) at $T = 298.2\ K$

Water-rich phase			Solvent-rich phase		
x_1^I	x_2^I	x_3^I	x_1^{II}	x_2^{II}	x_3^{II}
0.9800	0.0092	0.0108	0.1801	0.0448	0.7751
0.9700	0.0181	0.0119	0.2185	0.0612	0.7203
0.9600	0.0265	0.0135	0.2606	0.0757	0.6637
0.9500	0.0347	0.0153	0.3063	0.0892	0.6045
0.9400	0.0427	0.0173	0.3538	0.1012	0.5450
0.9300	0.0504	0.0196	0.4007	0.1111	0.4881
0.9200	0.0575	0.0225	0.4444	0.1187	0.4369
0.9100	0.0642	0.0258	0.4837	0.1241	0.3922
0.9000	0.0702	0.0298	0.5188	0.1274	0.3538
0.8900	0.0756	0.0344	0.5501	0.1290	0.3209
0.8800	0.0821	0.0379	0.5765	0.1324	0.2910
0.8700	0.0867	0.0433	0.6017	0.1324	0.2659
0.8600	0.0911	0.0489	0.6242	0.1323	0.2435
0.8500	0.0951	0.0549	0.6446	0.1316	0.2237
0.8400	0.0989	0.0611	0.6632	0.1307	0.2061
0.8300	0.1023	0.0677	0.6802	0.1296	0.1902
0.8200	0.1054	0.0746	0.6959	0.1282	0.1759
0.8100	0.1083	0.0817	0.7103	0.1267	0.1630
0.8000	0.1110	0.0890	0.7236	0.1252	0.1511
0.7900	0.1134	0.0966	0.7360	0.1235	0.1405
0.7800	0.1156	0.1044	0.7475	0.1217	0.1308

Table 4 Model tie-line data for water (1) + propionic acid (2) + dimethyl maleate (3) at $T = 298.2\ K$

Water-rich phase			Solvent-rich phase		
x_1^I	x_2^I	x_3^I	x_1^{II}	x_2^{II}	x_3^{II}
0.9800	0.0084	0.0116	0.2279	0.0473	0.7249
0.9700	0.0158	0.0142	0.2656	0.0579	0.6764
0.9600	0.0228	0.0172	0.2995	0.0751	0.6254
0.9500	0.0298	0.0202	0.3298	0.0800	0.5902
0.9400	0.0366	0.0234	0.3583	0.0890	0.5527
0.9300	0.0431	0.0269	0.3850	0.0971	0.5179
0.9200	0.0494	0.0306	0.4104	0.1043	0.4853
0.9100	0.0556	0.0344	0.4347	0.1107	0.4547
0.9000	0.0613	0.0387	0.4845	0.1204	0.3950
0.8900	0.0671	0.0429	0.4845	0.1218	0.3937
0.8800	0.0725	0.0475	0.4941	0.1237	0.3822
0.8700	0.0777	0.0523	0.5153	0.1271	0.3576
0.8600	0.0827	0.0573	0.5346	0.1297	0.3357
0.8500	0.0873	0.0627	0.5527	0.1317	0.3156
0.8400	0.0918	0.0682	0.5700	0.1331	0.2969
0.8300	0.0960	0.0740	0.5864	0.1341	0.2794
0.8200	0.0999	0.0801	0.6022	0.1347	0.2631
0.8100	0.1036	0.0864	0.6172	0.1350	0.2478
0.8000	0.1071	0.0929	0.6316	0.1348	0.2336
0.7900	0.1103	0.0997	0.6453	0.1345	0.2202
0.7800	0.1134	0.1066	0.6584	0.1338	0.2077
0.7700	0.1161	0.1139	0.6709	0.1329	0.1962
0.7600	0.1187	0.1213	0.6829	0.1318	0.1852

Table 5 Experimental and model tie-line data for water (1) + carboxylic acid (2) + dimethyl maleate (3) at $T = 298.2\ K$

Water-rich phase						Solvent-rich phase					
x_1		x_2		x_3		x_1		x_2		x_3	
Exp.	Model	Exp.	Model	Exp.	Model	Exp.	Model	Exp.	Model	Exp.	Model
Water (1) + butyric acid (2) + dimethyl maleate (3) $RMSD = 9.28 \times 10^{-5}$											
0.9760	0.9760	0.0153	0.0153	0.0087	0.0087	0.1366	0.1366	0.0279	0.0279	0.8355	0.8355
0.9571	0.9571	0.0341	0.0341	0.0088	0.0088	0.1969	0.1969	0.0482	0.0482	0.7549	0.7549
0.9370	0.9370	0.0532	0.0532	0.0098	0.0098	0.2391	0.2391	0.0632	0.0632	0.6977	0.6977
0.8987	0.8987	0.0873	0.0873	0.0140	0.0140	0.2940	0.2940	0.0839	0.0839	0.6221	0.6221
0.8575	0.8575	0.1201	0.1201	0.0224	0.0224	0.3425	0.3425	0.1032	0.1032	0.5543	0.5543
0.8132	0.8132	0.1504	0.1505	0.0364	0.0363	0.3978	0.3981	0.1265	0.1267	0.4757	0.4753
Water (1) + acetic acid (2) + dimethyl maleate (3) $RMSD = 4.56 \times 10^{-5}$											
0.9783	0.9783	0.0104	0.0104	0.0113	0.0113	0.1501	0.1501	0.0433	0.0433	0.8066	0.8066
0.9585	0.9585	0.0265	0.0265	0.0150	0.0150	0.2621	0.2621	0.0855	0.0855	0.6524	0.6524
0.9357	0.9356	0.0432	0.0432	0.0212	0.0212	0.3987	0.3988	0.1157	0.1157	0.4856	0.4855
0.9048	0.9049	0.0620	0.0620	0.0331	0.0331	0.5849	0.5850	0.1246	0.1246	0.2904	0.2904
Water (1) + propionic acid (2) + dimethyl maleate (3) $RMSD = 3.3 \times 10^{-3}$											
0.9758	0.9758	0.0115	0.0115	0.0127	0.0127	0.2443	0.2443	0.0518	0.0518	0.7039	0.7038
0.9584	0.9584	0.0240	0.0238	0.0176	0.0178	0.3042	0.3068	0.0723	0.0836	0.6234	0.6096
0.9314	0.9314	0.0422	0.0422	0.0264	0.0264	0.3814	0.3813	0.0960	0.0960	0.5227	0.5226
0.8975	0.8975	0.0628	0.0628	0.0397	0.0397	0.4829	0.4836	0.1206	0.1207	0.3965	0.3956

Table 5 continued

| Water-rich phase						Solvent-rich phase					
x_1		x_2		x_3		x_1		x_2		x_3	
Exp.	Model	Exp.	Model	Exp.	Model	Exp.	Model	Exp.	Model	Exp.	Model
0.8492	0.8492	0.0877	0.0877	0.0631	0.0631	0.5542	0.5541	0.1318	0.1318	0.3140	0.3141
Water (1) + formic acid (2) + dimethyl maleate (3) $RMSD = 1.14 \times 10^{-4}$											
0.9707	0.9707	0.0175	0.0175	0.0118	0.0118	0.2157	0.2157	0.0601	0.0601	0.7243	0.7242
0.9519	0.9519	0.0332	0.0332	0.0149	0.0149	0.2975	0.2973	0.0867	0.0867	0.6158	0.6160
0.9165	0.9165	0.0599	0.0599	0.0236	0.0236	0.4587	0.4587	0.1208	0.1208	0.4205	0.4205
0.8830	0.8830	0.0808	0.0810	0.0362	0.0360	0.5683	0.5681	0.1325	0.1328	0.2992	0.2991
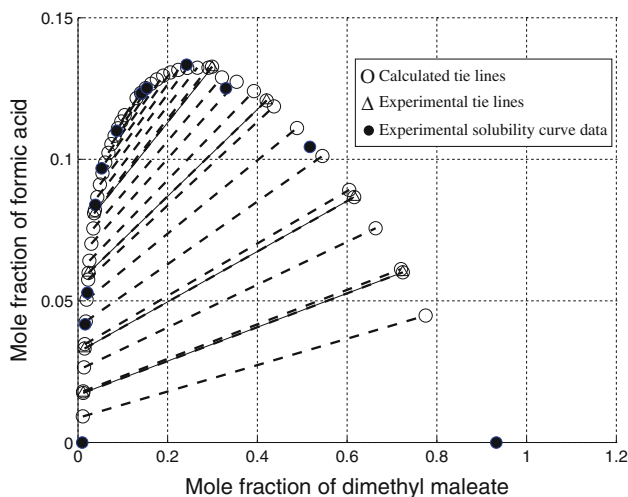
Fig. 4 Experimental and calculated LLE data for water–formic acid–dimethyl maleate ternary system
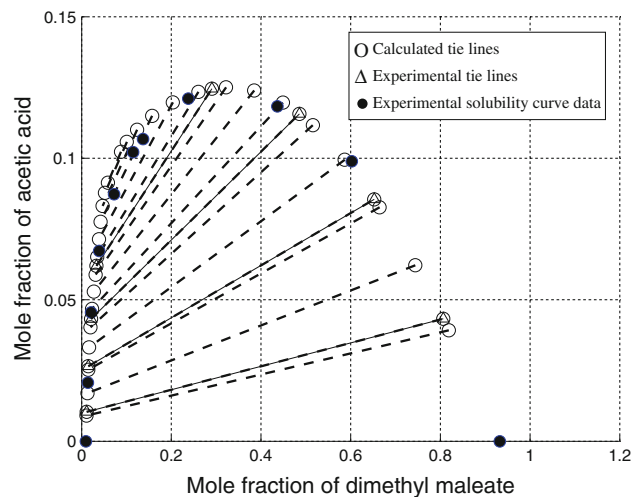

Fig. 5 Experimental and calculated LLE data for water–acetic acid–dimethyl maleate ternary system

Table 5 for experimental and corresponding calculated data. RMSD values are found varying between 4.56×10^{-5} and 3.3×10^{-3}. The correlated results and solubility curve data were plotted in Figs. 4, 5, 6, and 7 along with the experimental values.

7 Conclusion

In this paper, a new method for correlation of ternary LLE data is presented. The method is implemented by using a combined structure that uses genetic algorithm (GA)–trained neural network (NN). The LLE data were correlated using the NN-based activity coefficient models. The correlation with the NN model gives much better results than the NRTL and UNIQUAC equations for the studied systems [7]. It is apparent from the figures that the solubility curve and tie-line data agree well with the calculated data obtained from the proposed method. In this study water +
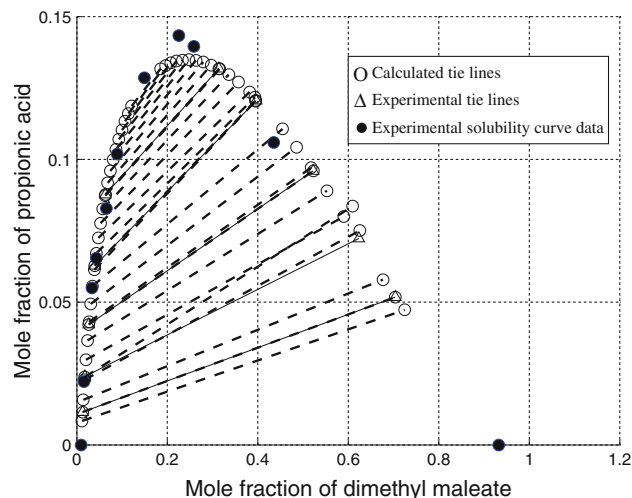

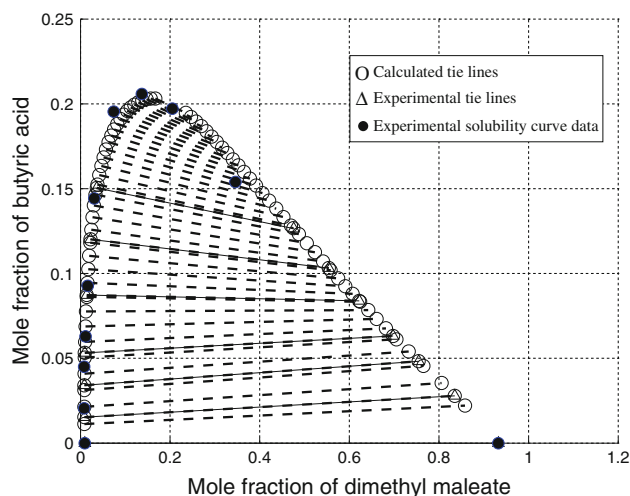**Fig. 6** Experimental and calculated LLE data for water–propionic acid–dimethyl maleate ternary system

Fig. 7 Experimental and calculated LLE data for water–butyric acid–dimethyl maleate ternary system

carboxylic acid + dimethyl maleate systems were evaluated, and of course, more systems will be processed in the studies of future. As shown in this work this method has potential to obtain tie-line and solubility curve data of LLE by using quite fewer experimental data.

References

1. Poling BE, Prausnitz JM, O'Connell JP (1998) Properties of gases and liquids, 5th edn, vol 8. McGraw-Hill, Boston, pp 1–204
2. Seader JD, Henley EJ (1998) Separation process principles, 1st edn. Wiley, New York, pp 31–89
3. Mohanty S (2005) Estimation of vapour liquid equilibria of binary systems, carbon dioxide ethyl caproate, ethyl caprylate and ethyl caprate using artificial neural networks. Fluid Phase Equilibria 235:92–98
4. Mohammadi AH, Martinez-Lopez JF, Richon D (2010) Determining phase diagrams of tetrahydrofuran+methane, carbondioxide or nitrogen clathrate hydrates using an artificial neural network algorithm. Chem Eng Sci 65:6059–6063
5. Nguyen VD, Tan RR, Brondial Y, Fuchino T (2007) Prediction of vapor-liquid equilibrium data for ternary systems using artificial neural networks. Fluid Phase Equilibria 254:188–197
6. Singh MK, Banerjee T, Khanna A (2005) Genetic algorithm to estimate interaction parameters of multicomponent systems for liquid-liquid equilibria. Comput Chem Eng 29:1712–1719
7. Özmen D (2008) Determination and correlation of liquid-liquid equilibria for the (water + carboxylic acid + dimethyl maleate) ternary systems at T = 298.2K. Fluid Phase Equilibria 269:12–18
8. Ghanadzadeh H, Ganji M, Fallahi S (2012) Mathematical model of liquid-liquid equilibrium for a ternary system using the GMDH-type neural network and genetic algorithm. Appl Math Model 36:4096–4105
9. Daubert TE (1987) Chemical engineering thermodynamics, 2nd edn, vol 7. McGraw-Hill, Singapore, pp 259–290
10. Treybal RE (1963) Liquid extraction, 2nd edn. McGraw Hill Book Company, New York pp 56–121
11. Prausnitz JM, Lichtenthaler RN, de Azevedo EG (1999) Molecular thermodynamics of fluid-phase equilibria, 3rd edn. Prentice Hall, New Jersey, pp 213–304
12. Prausnitz JM, Anderson TF, Grens EA, Eckert CA, Hsieh R, O'Connell JP (1980) Computer calculations for multicomponent vapor-liquid and liquid-liquid equilibria. Prentice-Hall, Englewood Cliffs, pp 39–81
13. Sorensen JM (1980) Correlation of liquid-liquid equilibrium data, PhD. thesis, Technical University of Denmark, Lyngby, Denmark
14. Haykin S (1999) Neural networks, 2nd edn. Prentice-Hall, New Jersey, pp 6–23
15. Şengül M, Özmen A, Yılmaz M (2003) A new synthesis algorithm for mixed lumped-distributed low-pass ladder networks via artificial neural networks. International conference on electronics and computer engineering, Bursa, Turkey.
16. Tang KS, Man KF, Kwong S, He Q (1996) Genetic algorithms and their applications. IEEE Signal Process Mag 13:22–37
17. Mitchell M (1999) An introduction to genetic algorithms, 5th edn. A Bradford Book The MIT Press, London, pp 7–12
18. Lagarias JC, Reeds JA, Wright MH, Wright PE (1998) Convergence properties of the nelder-mead simplex method in low dimensions. SIAM J Optim 9:112–147