J. of Supercritical Fluids 84 (2013) 80–88

Contents lists available at ScienceDirect

The Journal of Supercritical Fluids

journal homepage: www.elsevier.com/locate/supflu

Phase equilibria modeling of binary systems containing ethanol using optimal feedforward neural network

B. Vaferi[a,*], Y. Rahnama[b], P. Darvishi[c], A. Toorani[a], M. Lashkarbolooki[d]

[a] Department of Mechanic, Beyza Branch, Islamic Azad University, Fars, Iran
[b] School of Chemical and Petroleum Engineering, Shiraz University, Shiraz, Iran
[c] Chemical Engineering Department, School of Engineering, Yasouj University, Yasouj, Iran
[d] Islamic Azad University, Dashtestan Branch, Borazjan, Iran

ARTICLE INFO

Article history:
Received 28 April 2013
Received in revised form
26 September 2013
Accepted 27 September 2013

Keywords:
Binary system
Bubble point pressure
Vapor phase composition
Artificial neural network
Optimal configuration

ABSTRACT

A comprehensive understanding of vapor liquid equilibrium (VLE) data is one of the most important information for designing and modeling of process equipment. Because, it is not always possible to completely carry out experiments at all of the possible operational temperatures and pressures range, generalized thermodynamic models, e.g. equations of state are constructed for computing of required VLE data. In this work, artificial neural network (ANN) was used to derive predictive models of bubble point pressure and vapor phase composition of binary ethanol (C_2H_5OH) mixtures. In the neural network model, it is assumed that the considered VLE data depend on critical properties, acentric factor, normal boiling point, liquid phase composition of the solutes, and temperature. The proposed ANN model has been constructed and trained with VLE experimental data of nine different binary systems containing C_2H_5OH collected from various literatures. Optimal configuration of the ANN model has been determined using minimizing the average absolute relative deviation percent (%AARD), mean square errors (MSE) and the maximizing the correlation coefficient (R^2) between observed and predicted VLE data with the ANN model. By using this procedure a two-layer ANN model with twenty-three hidden neuron has been found as an optimal topology. The accuracy of our optimal two layers ANN model has been compared with the Peng–Robinson cubic equation combined with Wong–Sandler (WS) mixing rules including a Van Laar (VL) model for the excess Gibbs free energy. Comparison with available literatures data and Peng–Robinson equation of state confirm that the present ANN model is more accurate and superior than the other published works. The sensitivity errors analysis clarify that our ANN model could predict vapor phase composition and bubble point pressure of all of the nine binary ethanol systems with %AARD of 1.52% and 2.59% respectively. The study demonstrates that the neural network model is a good alternative method for the estimation of VLE properties of the binary system containing C_2H_5OH.

1. Introduction

The phase behavior of the binary mixtures containing ethanol has received great interests from academia and industry in the last few decades. Thermodynamic and phase behavior of ethanol + congeners (compounds which different from ethanol) mixtures is essential to design, simulation and control of related equipment units such as distillation in the alcoholic beverage production processes [1]. In addition, some of the congener substances constitute the aroma part of the distilled product and therefore their concentrations are important parameters [2–4]. While the concentration of congeners substances provides flavor and aroma characteristics to the final distilled spirit and also because their concentrations are regulated by law, the precise modeling of the vapor phase concentration of these materials have attracted popular attentions [5].

Widely used technique to predict VLE properties such as bubble point pressure and vapor phase concentration of the ethanol–congeners system is the combination of an equation of state (EOS) with a model for the excess Gibbs free energy. Faúndez et al. [3] analyzed vapor–liquid equilibrium in binary ethanol + congeners mixtures found in alcoholic distillation by using the Peng–Robinson EOS [6] and one of the most popular modern mixing rules, i.e. Wong–Sandler model [7,8]. They have used Van Laar model [3] for the Gibbs excess energy [8] in Wong–Sandler mixing rules. ANN model is another attractive technique which has been proposed to predict VLE properties [9]. The greatest advantage of an ANN is eliminating the complex equations, and replacing them with popular transfer functions [10]. Recently,

* Corresponding author. Tel.: +98 7116135904; fax: +98 7116473687.
E-mail addresses: behzad.vaferi@gmail.com, behzad.vaferi@yahoo.com (B. Vaferi).

0896-8446/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.supflu.2013.09.013

Lashkarbolooki et al. [9] utilized ANN model for prediction of solid solubilities in supercritical carbon dioxide and compared their ANN results with Peng–Robinson and Soave–Redlich–Kwong EOSs. They have highlighted the advantages of ANN over the combination of an EOS with different mixing rules for estimating the VLE properties [9]. Rohani et al. reported that ANN model can be considered as a suitable tool for estimation of phase equilibria of complex systems [10]. Mohanty designed an ANN model to simulate the phase behavior of binary systems including: carbon dioxide–ethyl caproate, ethyl caprylate and ethyl caprate in the temperature range of 308.2–328.2 K and pressure range from 1.6 to 9.2 MPa [11]. Mohanty compared the results of developed ANN model by Peng–Robinson and Soave–Redlich–Kwong equations of state [12]. In present work, a feed-forward back-propagation MLP network with only one hidden layer has been used to predict VLE properties of nine different binary ethanol mixtures. Predictions of ANN model were compared with available literature data and results obtained using combination of EOS and Wong–Sandler mixing rule [3,9].

2. Artificial neural network

Deriving reliable and precise analytical relation that can explain highly non-linear phenomena and correlate associated independent and dependent variables of realistic or synthetic processes are often difficult and sometimes impossible. Hence popularity of black box models which are generally based on artificial intelligence techniques such as ANNs, have been increased for simulating of complex process behavior in areas where precise analytical or semi-experimental correlation are unavailable. ANN systems are non-linear learning mathematical method which widely utilized for data processing, process analysis and control, fault diagnosis and pattern recognition [13]. These non-linear learning mathematical models were designed in the second half of the twentieth century by simulation of human brain procedures [13].

MLPNNs with totally feed-forward connections and back-propagation learning algorithm are among the most widely used ANNs models which are being extensively used in various fields of science and engineering up to now [14,15]. These networks are capable to correlate inputs and outputs of most non-linear multivariable phenomena with any complexity or in situation that no available relation at all. These networks composed of a large number of key processing elements that are connected together in a specified manner according to the type of the network.

Modeling based on ANN has been done in a way which does not require exact formulation of relations between input and output information or assumption about the parametric nature of the related parameters. Using of ANNs is increased because of their non-linearity, massive parallel connections, multiple input-output variables, non-requirement of assumption about the functional form of the model and also tolerant noisy data [13].

Each ANN models consist of a number of simple processing units that are connected together in a specified manner according to the type of the network. These processing units have been inspired from biological neurons, and are called neurons. Input signals always fed to input layer and then transfer to neurons in the hidden layers and output one, respectively. It can be said that the neurons in the output layer provide the results of MLPNN. The output of a neuron is computed from the Eq. (1):

$$n_j = f\left(\sum_{r=1}^{N} w_{jr}x_r + b_j\right) \tag{1}$$

As can be found from Eq. (1), the input signals to each neuron are weakened or strengthen through their multiplication to weight coefficients (w_{jr}). The biases (b_j) are activation thresholds that are added to the production of inputs (x_r) and their particular weight coefficients. The net output of each neuron passes through a function which is called activation or transfer function (f) of the neuron. Different types of transfer functions have been proposed for artificial neural networks such as linear, logarithmic sigmoid, hyperbolic tangent sigmoid, and radial basis transfer functions [13]. In the present study, the following function is utilized as the transfer functions in input as well as output layer:

$$f(x) = \frac{1}{1+\exp(-x)} \tag{2}$$

The correlation indicated by Eq. (2) is usually called log-sigmoid transfer functions. This transfer function compress input data into intervals [0 1]. The non-linearity, continuity and differentiability nature of aforementioned function allow the neuromorphic model to relate input and output data with any complexity. The differentiability of the transfer functions is important characteristic, which allows the gradient-based training algorithms to update the weights and biases.

2.1. Selection of optimum configuration

The most important issues in developing of any ANN model are specifying the optimal number of hidden layers and the number of neurons per each layer. Although back-propagation can be applied to networks with any number of layers it has been mathematically proven that any multivariable function with arbitrary discontinuities can be approximated to desired accuracy using the MLPNN with only one hidden layer provided non-linear transfer functions in its hidden units, i.e. sigmoid [16–19]. Cybenko substantiated his theory using the Hahn–Banach theorem [20] while the proof of Hornik et al. [16] is based on the Stone–Weierstrass theorem [20], and Funahashi [17] proved the same problem using an integral formula. Xiang et al. [21] proof is most elegant and simple, and is derived from a piecewise-linear approximation of the sigmoidal activation function. The appropriate number of hidden neurons for approximating a target function is not known in general and traditionally determines by trial and error procedure. Evaluation the optimum number of hidden units is complex task because there is considerable dependence on three issues: (1) complexity of correlation between input and output data being attacked using a neural network, (2) the number of available training and test data, and (3) the severity of noise imposed on the data sets [22]. A low number of neurons are not powerful enough to reach to the desired error and often lead to under-fitting, while a large number of neurons have a very expensive computation and may result in over-fitting [23].

In analogy to curve fitting, smaller networks that use fewer parameters usually have better generalization capability. During training an MLP, the optimal number of neurons in the hidden layers is unknown and is estimated usually by trial-and-error. Two strategies, namely, network pruning and network growing, are used to determine the size of the hidden layers [22].

To avoid over-fitting it is necessary that the flexibility of a developed ANN reduced [24]. Flexibility is directly related to the number of hidden layer and hidden neurons. Increasing the number of network parameters (weights) lead to increases the flexibility of ANN as well. In the other words, the larger number of weights causes the larger flexibility.

Some researchers have proposed penalties for choosing the number of nodes in the hidden layer [25,26]. Lippmann has proven that the numbers of MLP hidden nodes are often function of number of independent variables [27]. Mehrotra et al., suggest the optimization approach to estimate the number of hidden nodes required to solve a classification problem in ÷dimensional input space [28].