Journal of Molecular Liquids 197 (2014) 65–72

Contents lists available at ScienceDirect

Journal of Molecular Liquids

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

Predicting ionic liquid based aqueous biphasic systems with artificial neural networks

Shahla Shahriari [a,*], Shirin Shahriari [b]

[a] Department of Chemical Engineering, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran
[b] DMA—Department of Mathematics and Applications, CMAT—Centre of Mathematics, University of Minho, Portugal

ARTICLE INFO

Article history:
Received 14 November 2013
Received in revised form 12 April 2014
Accepted 28 April 2014
Available online 11 May 2014

Keywords:
Aqueous biphasic system
Ionic liquid
Entropy of hydration
Artificial neural network

ABSTRACT

The development of artificial neural networks to describe the formation of $[C_4mim][CF_3SO_3]$-based aqueous biphasic systems (ABSs) for a broad range of salts, composed of diverse combinations of cations and anions is reported here. Three different artificial neural networks were used for predicting the cation effect of the chloride- and acetate-based salts and the anion effects of sodium salts to induce the formation of ionic-liquid-based ABSs. For the development of ANN, the input data set was classified into three sets: seventy five percent of the total data were selected for training and twenty five percent of the data were selected for testing and validation. The results show that using the proposed neural network model a good agreement between the experimental data points and the predicted values was achieved. These results suggest that the ANN method may be a useful tool for predicting the formation of IL-based ABS.

1. Introduction

The purification of biological substances, such as enzymes, drugs, and antibodies, is one of the most expensive steps on their production and their economical viability is often dependent on the development of novel or improved downstream processes. During the last decade, ionic liquid-based-aqueous biphasic systems (ABSs) have become a promising alternative new technique for biomolecule analysis, separation and purification [1–6]. Considerable efforts have been focused on the use of ionic liquids (ILs) as substitutes for polymer-rich phase [2] and, recently also for salt-rich phase [7–11].

ILs are a kind of salt in which the ions are poorly coordinated, which results in solvents being liquid below 100 °C. The unique characteristics of these liquid salts make the ionic liquids interesting for industrial application, so that they are expected to revolutionize a number of areas as in the production of cellulose products, catalysis, extraction and separation processes [10,11]. The main physical and chemical properties of ILs are excellent solvation qualities, a wide liquidus range, nonvolatile, excellent chemical and thermal stabilities, good electrical conductivity and high ionic mobility [7]. A wide range of potential combinations between cations and anions allows a high degree of tunability of the IL properties, and the tailoring of, ILs for specific applications [10,11].

There is today a growing body of work on the field of ionic liquid-based aqueous biphasic systems [2], but most of these studies are based on experimental measurements of phase diagrams and partition coefficients for model molecules [10–16]. Given the large amount of available experimental data it is now required to develop modeling approaches that can provide a comprehensive vision of these systems, and present some predictive capabilities allowing for a faster and cheaper scanning of the most adequate aqueous biphasic system for a given separation. This scanning comprises two aspects, the definition of the phase diagram, and the study of the partition coefficients for a given macromolecule of interest.

There is, nowadays, an increasing trend towards the use of smart technology for solving complicated scientific problems in different industrial areas [17]. These systems learn general rules on the system or process behavior by using complex algorithms to relate them to the input experimental data. Artificial neural networks (ANNs) belong to a group of smart systems, which transfer the hidden knowledge beyond the data to the network structure by processing the experimental data. The biggest advantage of neural networks is their ability to model complex non-linear relationships regardless of previous assumptions. Artificial neural networks have features which distinguish them in some applications such as pattern and model recognition in complex and non-linear systems from other methods [18,19]. In the last ten years, the neural networks have attracted a lot of attention. These networks suggest an alternative way for computing akin to that used by the human brain. This approach is not a new one. In 1943, a neurophysiologist and a logician by the names of Warren McCulloch and Walter Pits produced the first artificial neuron [20]. During the sixties, people focused on the side of artificial intelligence and only in the eighties the real potential of neural networks had been grasped by the broader

* Corresponding author. Tel.: +98 2144802841.
E-mail address: shahla_shahriari@yahoo.com (S. Shahriari).

http://dx.doi.org/10.1016/j.molliq.2014.04.030
0167-7322/

community. Neural networks have unique features which distinguish them from other methods such as:

1- The ability to learn, which means the ability to adjust network weights with the use of training data.
2- The ability to generalize, which means that after network development with the use of training data and by adjusting network weights, the network is able to accept an input which is not taught, and provide an appropriate output.

In many fields where problems with complex relations between input and output are found, from biomedical diagnosis to the behavior of chemical systems, researchers are actively using ANN to tackle many of these problems [17]. The dramatic increase in the number of publications testifies the success of the application of neural networks to different fields in a wide variety of applications [21,22] from biomedical researches [23–25] to the simulation and control of processes in the food industry and agriculture [26–31], the modeling of lysine production [32], optimization of fed-batch bioreactor [33], the prediction of viscosity and density of ionic liquids [34,35] and prediction of thermodynamic properties [36] and the partitioning of biomolecules [37].

In a previous work, we have studied the formation of ABSs to gather a broader picture on the molecular mechanisms that govern the formation of these systems [38]. The ability of a broad range of salts, composed of diverse combinations of cations and anions, to induce the formation of $[C_4mim][CF_3SO_3]$-based ABSs was evaluated by means of the determination of the corresponding ternary phase diagrams at 298 K and atmospheric pressure. With the large body of data measured in this previous study, it was possible to establish a scale on the salt cation and anion abilities to induce the formation of ionic-liquid-based ABSs, which follows the Hofmeister series, and to show that the molar entropy of hydration of the salt ions is the driving force for aqueous two-phase system formation [38].

This study aims at examining the power of neural network models in forecasting the formation of ionic-liquid-based ABSs. The formation of IL-based ABSs can be evaluated in two dimensions:

1- The ability of a range of salts, composed of diverse combinations of cations and anions, to induce the formation of ABSs based on a selected ionic liquid;
2- The ability of a range of ionic liquids, composed of diverse combinations of cations and anions, to undergo the formation of ABSs with a selected salt.

This study presents the development of an artificial neural network to describe the formation of $[C_4mim][CF_3SO_3]$-based ABSs for a broad range of salts, composed of diverse combinations of cations and anions. If successful, a future development of this work will expand it towards a simultaneous description of the two dimensions (ILs and salts) in the ABS formation.

2. The artificial neural network

2.1. Structure of ANN

The motivation behind the studies of neural networks is originated from the fact that human brain computes in a totally different manner. It is an advanced complex, nonlinear and parallel processing system containing billions of structural constituents known as neurons. Since the brain has extensive ability to organize these neurons to process information in parallel, it can perform specific computation tasks much faster than the fastest computers available today [39].

The ANN is a powerful tool for describing complex input/output relationships and can also be used to mapping relationships from inputs to outputs of these kinds of black boxes. The true power and benefit of neural networks lies in their capability to represent linear and non-linear relationships and in their ability to learn these relationships directly from the data being modeled [40]. They also have a considerable ability to draw meaning from vague or complicated data and can be used for detecting highly complex patterns which are too complicated to be extracted by humans or other computer techniques. A trained neural network can be considered as an "expert" for analyzing the given information. Then by using this expert, the predictions for new systems of interest can be achieved. The ANN models have advantages such as: a) Adaptive learning: ANN are capable to learn how to do tasks on the basis of the data given for training. Therefore, there is no need for extraordinary feats of programming; b) Self-organization: An ability to produce its own organization and presentation of the receiving information during learning time; c) Real time operation: ANN has the capability to perform the computations in parallel, and particular hardware devices are being designed and made for taking advantage of this ability; d) Noise tolerant: ANN can cope with situations where normal symbolic systems would have difficulty; e) Indeed, ANN has the ability to do anything a symbolic/logic system can do [39,40].

The ANN has included one input, one output and one or more hidden layers. Each layer contains several neurons which are connected to each other by two coefficients called weights (w) and biases (b). Data processing is carried out in neurons, and the signals are transmitted among neurons by their connections. Each neuron imposes an attractivation function on its inputs which is the weighted sum of input signals so that it produces its output signal.

An ANN determines an empirical relationship between the inputs and the outputs of a given system where the inputs of the systems are the independent variables and the outputs are the dependent variables. The output of a neuron is determined according to the following equation:

$$n_j = \sum_{i=1}^{N} f\left[\left(w_{jr} I_r\right) + b_j\right] \tag{1}$$

where I_r is the input, w_{jr} is the weight parameter between neurons in the two layers, j is the hidden neurons, the term b_j relates to the bias of neuron j of the hidden layer and f is the transfer function of neurons [41].

The arrangement of neurons in layers and patterns of connections within and between the layers is called network architecture. The basic different classes of network architectures have been suggested in the literature. One of the most used networks is called multi-layer feed forward network (MLFFN) that is built on the set of interconnected neurons organized in layers corresponding to the input, hidden and output layers. This type is used to build a non-linear correlation between input and output parameters.

The input layer is fed with the input signals and these signals are then delivered to the hidden layer(s) where the processing task is performed. A neuromorphic model predicts the final values of the desired properties which are called the network's output and are supplied by the neurons in the output layer. The number of independent and dependent variables defines the number of neurons for the input and output layers respectively [42]. Finally, the last hidden layer delivers the information to the output layer which sends the results to an external source. This type of network has the ability to estimate most types of functions without considering how complicated they are [43]. During the training algorithm, input data are sent to the input layer of the network and the network's synaptic weights and biases are adjusted by using the network error as a criterion which is the difference between the results from the output layer and the desired outputs. The initial values of all synaptic weights and biases are chosen randomly and then, an optimization algorithm is used to train the network (i.e., its synaptic weights are adjusted) in order to emulate the input/output mapping [42,44]. Different kinds of transfer functions have been presented for ANN such as linear, logarithmic sigmoid, hyperbolic tangent sigmoid and