Journal of Molecular Liquids 211 (2015) 534–552

Contents lists available at ScienceDirect

Journal of Molecular Liquids

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

Prediction of the binary surface tension of mixtures containing ionic liquids using Support Vector Machine algorithms

Mohammad Hashemkhani [a], Reza Soleimani [b,*], Hossein Fazeli [c], Moonyong Lee [d], Alireza Bahadori [e], Mahsa Tavalaeian [f]

[a] Ahwaz Faculty of Petroleum Engineering, Petroleum University of Technology (PUT), Ahwaz, Iran
[b] Young Researchers and Elite Club, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran
[c] Department of Geosciences, University of Oslo, Oslo, Norway
[d] School of Chemical Engineering, Yeungnam University, Gyeungsan, Republic of Korea
[e] Southern Cross University, School of Environment, Science and Engineering, Lismore, NSW, Australia
[f] Department of Physics, University of Zanjan, Zanjan, Iran

ARTICLE INFO

Article history:
Received 1 February 2015
Received in revised form 10 June 2015
Accepted 14 July 2015
Available online 7 August 2015

Keywords:
Ionic liquids
Surface tension
Binary mixtures
Prediction
Support Vector Machine

ABSTRACT

The surface tension of pure ionic liquids (ILs) and their mixtures with other compounds play a key role in the design and development of many industrial processes. Therefore, its modeling is extremely important from an industrial point of view. This study examined the capability and feasibility of three intelligence algorithms for predicting the surface tension of binary systems containing ILs. To construct and test the models, 748 data points corresponding to the experimental surface tension values of binary mixtures containing ILs were extracted from the literature. The surface tension was between 0.0157 and 0.07185 $N \cdot m^{-1}$. The absolute temperature (T), mole fraction and molecular weight of the IL components (x_{IL} and Mw_{IL}) and the density of the IL components (ρ_{IL}) together with the boiling point ($Tb_{non\text{-}IL}$) and molecular weight ($Mw_{non\text{-}IL}$) of the non-IL component were considered as model input variables to differentiate between the various compounds involved in binary systems. A comparison of the experimental data and predicted values using all three methods (in terms of statistical parameters) showed good agreement; however, the CSA-LSSVM prediction was better than the other two approaches.

1. Introduction

Ionic liquids (ILs) are a novel class of purely ionic, salt-like materials comprised of bulky asymmetric organic cations and organic or inorganic anions that are unable to form an ordered crystal and thus remain in the liquid state at or close to room temperature [1].

ILs have attracted considerable interest in academia over the past few years and their introduction to industrial applications is ongoing [2,3]. The continually increasing number of publications and patents [4–8] on ILs in recent years provides a snapshot of the research and use of these fluids. The growing interest can be attributed to the exceptional features of these most versatile classes of materials, such as negligible vapor pressures, excellent solvating capacity for both organic and inorganic compounds, high thermal and chemical stability, high heat capacity, lack of inflammability, high thermal and electrical conductivity, and above all, their properties can be fine-tuned by judicious selection of the cation, anion and substituents [9], making them "tailor-made materials" that can be adjusted to satisfy the requirements of a given purpose.

The thermophysical properties require examination because it is important to screen the possibility of a transformation of ILs from the laboratory level to large-scale industrial applications. The identification, development and accomplishment of dependable and cost-effective process designs for industrial applications of ILs and their related mixtures rest largely upon a comprehensive study of their thermophysical properties, such as surface tension, density and viscosity [4].

Determining thermophysical properties of ILs accurately cannot be accomplished exclusively by experimental measurements due to the vast number of ILs possible. In addition, most industrial applications of ILs take place in mixtures, causing limitless combinations of mixtures whose thermophysical properties may be of interest [10]. Furthermore, the experimental measurements are intrinsically time-consuming and expensive with probable non-negligible uncertainties [11]. Therefore, it is important to develop predictive computational tools to estimate precisely the thermophysical properties of various systems containing ILs.

Surface tension is one of most important thermophysical properties providing information on the surface, intermolecular interactions and forces in a liquid, and has a considerable influence on the transfer of mass and energy across the interface [12]. Therefore, surface tension has a determining influence on various industrial processes, e.g., separation, extraction, adsorption, distillation, and performance of biological

* Corresponding author.
E-mail address: soleimanire@gmail.com (R. Soleimani).

http://dx.doi.org/10.1016/j.molliq.2015.07.038
0167-7322/

membranes [12]. Thorough descriptions of the importance of the surface tension of ILs are reported elsewhere [13].

Considerable efforts have been made to evaluate the surface tension of pure compounds using a variety of methods. A comprehensive review of these methods has been presented [11,14–16]. On the other hand, the literature reviews conducted by researchers [13,17] suggest that there are a few models for predicting the surface tension of ILs. Gharagheizi et al. [17] commented on the disadvantage of methods used by various researchers for modeling the surface tension of ILs. Generally, most methods have been already used on a desired system and not on a set of systems, which is the focus of this study. Indeed, they are not well generalized to predict the surface tension of a wide range of ILs. In addition, accurate knowledge of the structure of the ILs is an absolute necessity of group contribution methods (which was used by Gharagheizi et al. [17]), which in some cases, are complex and require considerable calculation time. In addition, the group parameters for various ILs functional groups have not been determined so these methods have limitations [18]. Therefore, it is important to study the capabilities of alternative techniques for the same purpose. In general, the salient features of the modeling techniques are their flexibility to model multiple mechanisms of actions, competence to handle high dimensional data and a high level of predictive accuracy [19].

Computational intelligence schemes provide effective and convenient predictive tools for modeling highly complex and non-linear systems. Artificial neural network (ANN) and evolutionary algorithms are the most well-known components of the so-called soft computational intelligence and have attracted considerable research attention in many fields in recent years. Several applications and impressive descriptions of ANN have been published [20,21]. Over the past few years, different approaches of ANNs have been applied successfully to analyze a wide range of properties [22], particularly when the focus is placed on predicting the properties of pure ILs as well as binary and ternary of mixtures containing ILs, including the ternary viscosity of mixtures [23], ternary electrical conductivity [24], binary heat capacity of mixtures [25], bubble points of ternary systems [10], density for binary mixtures of methanol + ILs [26], and to estimate the thermal conductivity of ILs [18]. Lazzús [27] reported the successful application of artificial neural networks to correlate a total of 2410 data points of density at several temperatures and pressures (ρ–T–P), corresponding to 250 ionic liquids. Torrecilla et al. [28] presented an optimized ANN model for predicting the melting point of a group of 97 imidazolium salts with various anions. Bini et al. [29] reported that a recursive neural network (RNN) was an applicable tool for predicting the melting points of several pyridinium based ionic liquids (ILs). Yiqing et al. [30] used the ANN model to predict the compositional viscosity of binary mixtures of room temperature ILs [C_n-mim] [NTf2] with n = 4, 6, 8, and 10 in methanol and ethanol over the entire range of molar fractions at temperatures ranging from T = 293.0–328.0 K. Eslamimanesh et al. [31] applied the ANN procedure to represent the solubility of supercritical CO_2 in 24 mostly used ionic liquids. Note that the pseudo critical properties of ionic liquids employed in their studies were estimated by group contribution method. Torrecilla et al. [32], in addition to Multiple Linear Regression (MLR) and Multiple Quadratic Regression (MQR), explored Radial Basis Network (RB), and MultiLayer Perceptron (MLP) Neural Network models for predicting the CO_2 solubility in 1-n-ethyl-3-methylimidazolium hexafluorophosphate, 1-n-hexyl-3-methylimidazolium hexafluorophosphate, 1-n-butyl-3-methylimidazolium tetrafluoroborate, 1-n-hexyl-3-methylimidazolium tetrafluoroborate, and 1-n-octyl-3-methylimidazolium tetrafluoroborate ILs under sub and supercritical conditions. Safamirzaei and Modarress [33] proposed a modeling method based on neural network technique and molecular properties to model the solubility of carbon dioxide, carbon monoxide, argon, oxygen, nitrogen, methane, and ethane in 1-butyl-3-methylimidazolium tetrafluoroborate, and they also used an ANN to correlate and predict Henry's law constants of gases in [bmim][PF_6] at low pressures [34]. The ANN-based Particle Swarm Optimization (PSO) method was used successfully to estimate the H_2S solubility in different ILs [35]. In addition, Genetic Expression Programming (GEP), as a new generation of evolutionary algorithms, was used effectively to predict the H_2S solubility in ILs [36]. Interested readers should refer to the literature [37–51], which provides valuable references on the application of different intelligence modeling approaches in the field of ILs.

Recently, specific attention has been dedicated to the Support Vector Machine (SVM) as a relatively novel intelligence method, which was first introduced by Vapnik [52]. Owing to its virtues and outstanding generalization performance over other methods, the SVM has attracted attention and been applied widely in pattern recognition, classification and regression problems [52–56]. SVM has a solid theoretical foundation based on the statistical learning theory [57]. The structural risk minimization (SRM) principle is employed in SVMs instead of the traditional empirical risk minimization principle (ERM) used in classical methods (such as conventional neural network), which makes SVMs quite resistant to the under-fitting and over-fitting problems [58,59]. In addition, because of their specific formulation, sparse solutions can be obtained, and both linear and nonlinear regression can be carried out [53]. A more advantageous SVM strategy over the traditional methods based on the ANNs can be found in the literature [60,61].

A least squares formulation of SVM, called LSSVM [62,63] was recently proposed to simplify the original formulation while retaining its many "good" properties. LSSVM reduces the complexity of calculation by converting a quadratic program into linear equality sets [64], resulting in an easier-to-implement method with a shorter computing time [65].

The application aspects of SVMs have been reported elsewhere [66]. A number of excellent introductions of SVM have been published [54, 63,67–70], and the theory of LSSVM was also described clearly by Suykens and Vandewalle [71] and Suykens et al. [63], Niazi et al. [72]. In addition, the application of LSSVM in quantification and classification was summarized by Niazi et al. [72]. Specifically, the results obtained from modeling the H_2S solubility in 11 different ILs by LSSVM based Genetic Algorithm (GA) [73], indicates that this method can be used as a potential candidate in the field of ILs.

A review of the literature indicated that aside from Oliveira et al. [74], who applied the soft-SAFT equation of state coupled with the density gradient theory (DGT) for modeling surface tension of binary mixtures of ILs, there are no publications concerning modeling of the surface tension of multicomponent systems including ILs. In the present study, SVM and its new variant, Least Square Support Vector Machine (LSSVM), were applied to estimate the surface tension of binary mixtures containing ILs. To the best of the authors' knowledge, there are no reports of the application of artificial intelligence schemes to the prediction of the surface tension of multicomponent systems containing ILs.

2. Data acquisition

The experimental data used for modeling consisted of 748 binary surface tension (*ST*) data points at atmospheric pressure assembled from The NIST Standard Reference Database [75,76]. The database contained binary surface tension data for 15 unique ILs. Table 1 lists the structures of the ILs investigated [75,76]. Table 2 presents the studied binary systems and the range of measured surface tensions (STs) in each system accompanied by the method used for measuring the *ST* [77–85]. After identifying and assembling the data set, the subsequent step was the selection of input variables, which are the model's independent variables. In this regard, the operational temperature (T), IL component compositions (x_{IL}), molecular weight of the IL components (Mw_{IL}), and the density of IL components (ρ_{IL}) along with the boiling point ($Tb_{non\text{-}IL}$) and molecular weight ($Mw_{non\text{-}IL}$) of non-IL component were introduced as entrance parameters. Table 3 presents the Mw_{IL} and ρ_{IL}; in addition, the $Tb_{non\text{-}IL}$ and $Mw_{non\text{-}IL}$ are listed in Table 4 [86–95]. The functionality of the binary surface tension of ILs can be