# Prediction of surface tension of ionic liquids by molecular approach

Seyyed Alireza Mirkhani [a,*], Farhad Gharagheizi [a,*], Nasrin Farahani [b], Kaniki Tumba [c]

[a] Department of Chemical Engineering, Buinzahra Branch, Islamic Azad University, Buinzahra, Iran
[b] Department of Chemistry, Buinzahra Branch, Islamic Azad University, Buinzahra, Iran
[c] Department of Chemical Engineering, Mangosuthu University of Technology, Durban, South Africa

## ARTICLE INFO

## ABSTRACT

Originally, Quantitative Structure Property Relationship (QSPR) models for the surface tension of ionic liquids are developed based on molecular descriptors. A large data set of 930 experimental surface tension data points for 48 ionic liquids is applied to derive the model. Seven descriptors are selected by genetic function approximation to relate the surface tension of ionic liquids to their corresponding anions and cation structures. To capture the nonlinear nature of surface tension, a model based on Least-Squared Supported Vector Machine (LSSVM) is also developed. The derived models are authenticated with several statistical validation techniques.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Recently, ionic liquids (ILs) as the new generation of conductive materials find their way in the innovative industrial and chemical applications. The term ionic liquid or more specifically room temperature ionic liquid (RTIL) refers to the salts entirely composed of ions and have a melting point below the normal boiling point of water. Same as other salts, ionic liquids possess very negligible vapour pressures even at conditions well above room temperature [1]. In addition, low-toxicity as well as non-volatility is the other novel characteristics of ILs. By possessing aforementioned qualities, ionic liquids are promising candidates to supersede convenient organic solvents in industrial applications.

Ionic liquids consisted of low-symmetry, large and unreactive cation such as phosphorus or sulphur containing ring and an anion that largely controls its physical and chemical properties. Altering in the cation and anion combinations permits the physical, chemical and biological properties of ionic liquids to be tailored for specific applications, as largely manifested by the task-specific ionic liquids (TSILs).

One of the recent applications of ILs is their employment in the processes such as extraction and multiphasic homogeneous catalytic reactions which are mainly governed by interfacial phenomena. The mentioned type of reactions occurs in the two-phase systems; one phase contains the reactant and products and the other, the immiscible one, act as the catalyst solvent. Such processes occur at the interface between the IL and the overlying aqueous or organic phase, and are dependent on the accessibility of the material to the surface and the transfer of material across the interface. More exhaustive studies of the surface-related properties are required to enhance our insight into the mechanisms behind these processes.

The available surface tension data of ionic liquids are very limited in comparison of possible ionic liquids ($\approx 10^{15}$). To "tailor" an ionic liquid with the desired properties, it is important to have a rational estimation of its properties e.g. surface tension prior to the synthesis in the absence of the experimental data. For this purpose, prediction methods which provide accurate estimations of desired properties are essential.

Models based on parachors, group contribution methods or corresponding state theories (CST) are widely applied for the prediction of the surface tension of the ionic liquids [2].

The foundation of parachor-based models are an empirical formula originally proposed by MacLeod [3], which relates density to the surface tension via temperature-independent Eq. (1):

$$\sigma^{\frac{1}{4}} = K\,\rho \tag{1}$$

Sugden [4] modified the original formula by multiplying each side of the expression by the molecular weight, $M_w$, to give a constant $K.M_w$ labelled as parachor, $P_{ch}$:

$$P_{ch} = KM_W = \frac{M_W \cdot \sigma^{\frac{1}{4}}}{\rho}. \tag{2}$$

Sugden [4] also proposed that the parachor of a compound is an additive property which can be described by the sum of its parachor contribution.

* Corresponding authors. Fax: +98 21 77 92 65 80.
E-mail addresses: seyyed.alireza.mirkhani@gmail.com (S.A. Mirkhani), fghara@gmail.com (F. Gharagheizi).

Knotts et al. [5] compiled a vast amount of available physical DIPPR data to derive a group contribution correlation for the estimation of parachors. For their model, deviations of 8% for multifunctional compounds were obtained with maximum deviation of 34%.

Deetlefs et al. [6] were the first who applied Knotts et al. [5] parachors for ionic liquids. They found that the differences between estimated parachor values based on non-ionic solvents and experimental ionic liquids' parachors are rather small. Although their calculations are based on the very limited data, they postulated that the QSPR correlation based on neutral species could be used for ionic liquids.

Gardas and Coutinho [7] are the first who successfully tackled the estimation of the surface tension of ionic liquids by applying Knotts et al. [5] parachor approach. By compiling 361 data points of 38 imidazolium based ionic liquids with thirteen anions, they derived a model with the overall deviation of 5.75% and maximum deviation less than 16%. Another fascinating outcome of their model is the successful application of Knotts' correlation, which is basically derived for neutral solvents without regard Coulombic interactions, for ionic liquids.

Finally, Mousazadeh and Faramarzi [8] applied CST methodology to derive a model for surface tension of ionic liquids. Since the CST correlations are primarily based on the critical properties, which are not available for ionic liquids, they chose to use the melting ($T_m$) and boiling points of ionic liquids ($T_b$) along with the surface tension at the melting point ($\sigma_m$) to define their corresponding states correlation.

$$\sigma = 0.819 \left( \frac{T_b - T}{T_b - T_m} \right) \sigma_m + 0.5 \left( \frac{T}{T_b} \right) \sigma_m. \tag{3}$$

The deviations reported for the surface tension of the 30 ionic liquids used in the development of the correlation represent only 3.0% while predictions for 4 ionic liquids not used in the correlation development are of 6.5%. Since many ionic liquids don't have a reliable melting point and determination of their boiling points is as elusive as their critical temperatures, the application of the proposed CST model for practical estimation of surface tension of ionic liquids is extremely limited.

In this communication, the predictive model derived from molecular descriptor is presented to estimate the surface tension of the studied ionic liquids.

Gharegheizi et al. [9] proposed a group contribution model to predict surface tensions of 51 ionic liquids based on 920 experimental data. 19 chemical substructures (12 for anions and 7 for cation) were considered to develop a group contribution model. The model generates acceptable results in terms of $R^2 = 0.919$ and Average absolute relative deviation (ARD) = 3.61. In comparison with our new QSPR model, GC model has simpler terms which are easily accessible from the chemical structure. In term of number of variables, the new QSPR model with 7 descriptors is superior than previous GC model with 19 parameters.

## 2. Methodology

### 2.1. Data preparation

An extensive literature review is conducted to extract 930 experimental surface tension data points belonging to 48 ionic liquids [10–29]. The collected data covers a wide range of surface tension (0.02315–0.06399 N.m$^{-1}$) as well as temperature (268.3–743.7 K). Table 1 describes the present ionic liquids with the range of their surface tension as well as temperature.

The structures of 12 anions and 19 cations present in our study are tabulated as Tables 2 and 3, respectively.

### 2.2. Training and test set selection

Typically, in QSPR modelling, the selected experimental data set is split into two subsets: the subset which involved in model derivation is called training set. The other one (test set) is used to assess the learning ability of the model from training set to produce reliable results for absent compounds. In this study, K-means clustering is applied to select training and test sets. *K-means* clustering is a method of cluster analysis which aims to partition $n$ observations into $k$ clusters in which each observation belongs to the cluster with the nearest mean. As the rule of thumb, 20% of collecting data was retained to test the model and remain was applied for model derivation [30]. For LSSVM model derivation, 80%–10%–10% of data points split into Training-Validation-Test sets respectively. This selection like the previous one is performed by k-mean clustering.

### 2.3. Calculation of descriptors

The aim of this study is to relate the surface tension of ionic liquids to their ionic constituents' structures via anion- and cation-based descriptors. For this purpose, anion and cation descriptors are separately calculated for each ionic liquid. This approach is successful to correlate the studied physical property with the structure of both anion and cation. However, it discarded anion/cation interactions.

In order to optimize the 3D chemical structure of each cation and anion, the Dreiding Force field as explained by Chemaxon's JChem was employed [31].

About 3000 descriptors from 22 diverse classes of descriptors are calculated by Dragon software [32]. These 22 classes of descriptors are: Constitutional descriptors, Topological indices, Walk and path counts, Connectivity indices, Information indices, 2D autocorrelations, Burden Eigen values, Edge-adjacency indices, Functional group counts, Atom-centred fragments, Molecular properties, topological charge indices, Eigenvalue-based indices, Randic molecular profiles, geometrical descriptors, RDF descriptors, 3D-MORSE descriptors, WHIM descriptors, GETAWAY descriptors, charge descriptors, 2D binary fingerprint, and 2D frequency finger print.

After the completion of descriptor calculation, those couldn't be calculated for several anions or cations are excluded completely from the list. Next, the pair correlations for each binary group of descriptors (all anions and cations descriptors) are calculated. For binary groups with the pair correlation greater than 0.9, one of the descriptors is omitted randomly.

### 2.4. Diversity test

To ensure the diversity of ionic liquids present in both training and test sets the diversity test is conducted in this study. There are many approaches to measure diversity, owing to its diverse definitions. In this study, the Euclidean distance approach is applied to measure the diversity of present ionic liquids. In this approach, each IL ($X_i$) is described by a vector of corresponding molecular descriptors incorporating both anion and cation descriptors ($x_{im}$) as its elements where m is the number of all calculated descriptors.

$$X_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{im}). \tag{4}$$

A distance of two different ionic liquids is defined as follows:

$$d_{ij} = \|X_i - X_j\| = \sqrt{\sum_{k=1}^{m} \left( x_{ik} - x_{jk} \right)^2}. \tag{5}$$

Next, the mean distance of one sample to the remaining ones is calculated as follows:

$$\overline{d_i} = \frac{\sum_{j=1}^{n} d_{i,j}}{n-1}. \tag{6}$$