Selection of optimum ionic liquid solvents for flavonoid and phenolic acids extraction

N R A Rahman¹,², N A Yunus¹,² and A A Mustaffa¹,²,³

¹ Faculty of Chemical and Energy Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.
² Process Systems Engineering Centre (PROSPECT), Research Institute for Sustainable Environment (RISE), Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.

Email: azizulazri@utm.my (A A Mustaffa)

Abstract. Phytochemicals are important in improving human health with their functions as antioxidants, antimicrobials and anticancer agents. However, the quality of phytochemicals extract relies on the efficiency of extraction process. Ionic liquids (ILs) have become a research phenomenal as extraction solvent due to their unique properties such as unlimited range of ILs, non-volatile, strongly solvating and may become either polarity. In phytochemical extraction, the determination of the best solvent that can extract highest yield of solute (phytochemical) is very important. Therefore, this study is conducted to determine the best IL solvent to extract flavonoids and phenolic acids through a property prediction modeling approach. ILs were selected from the imidazolium-based anion for alkyl chains ranging from ethyl > octyl and cations consisting of Br, Cl, [PF₆], [BF₄], [H₂PO₄], [SO₄], [CF₃SO₃], [TF₂N] and [HSO₄]. This work are divided into several stages. In Stage 1, a Microsoft Excel-based database containing available solubility parameter values of phytochemicals and ILs including its prediction models and their parameters has been established. The database also includes available solubility data of phytochemicals in IL, and activity coefficient models, for solid-liquid phase equilibrium (SLE) calculations. In Stage 2, the solubility parameter values of the flavonoids (e.g. kaempferol, quercetin and myricetin) and phenolic acids (e.g. gallic acid and caffeic acid) are determined either directly from database or predicted using Stefanis and Marrero-Gani group contribution model for the phytochemicals. A cation-anion contribution model is used for IL. In Stage 3, the amount of phytochemicals extracted can be determined by using SLE relationship involving UNIFAC-IL model. For missing parameters (UNIFAC-IL), they are regressed using available solubility data. Finally, in Stage 4, the solvent candidates are ranked and five ILs, ([OMIM] [TF₂N], [HeMIM] [TF₂N], [HMIM] [TF₂N], [HeMIM] [CF₃SO₃] and [HMIM] [CF₃SO₃]) were identified and selected.

1. Introduction
Herbal plants have played a major role in the treatment of human ailments since ages ago. It also known as main resources for food, flavouring, medicines and even function as fragrances. Herbs can synthesize wide range of bioactive compounds known as phytochemicals that can act as important biological function such as antibacterial, antiviral, anticancer and antifungal. Due to their high efficacy in treatment, natural herbs gain acceptances all over the world where 88 percent of human population now depends on traditional herbal medicine as their main daily health care [1]. The high demand of natural
herbs worldwide have encouraged alternatives and improvements to produce high quality herbal products.

*Labisia pumila* (*L. pumila*) or locally known as *Kacip Fatimah* is the most common herb used in Malaysian community for a variety of illnesses and also used as health supplements. *L. pumila* is well-acknowledged for their antioxidant and anticancer activity. Phenolic acids and flavonoids are the main phytochemicals of *L. pumila* that are responsible for variety of pharmacological activities of the herb [2]. Typically, phytochemicals can be identified and characterized from various plant parts such as leaves, stem, flower and fruits showed that the leaves of *L. pumila* exhibit higher antioxidant activities and total saponins amounts compared to the plant roots and stems [3][4].

The qualitative and quantitative of phytochemicals relies on the proper selection of the extraction method [5]. Commonly, solvents are used to extract phytochemical from herbs. Therefore, solvent selection is highly important to ensure the effectiveness of the solvent to extract phytochemicals from herbs [6]. Different solvents have different capacities of extraction depending on the solubility and polarity of the phytochemicals [7].

Current herbal extraction processes used organic solvents such as ethanol, methanol, chloroform and acetone. However, these conventional solvents are volatile, toxic and the selectivity of extraction is low. Therefore, ionic liquid (IL) has been considered as an alternative solvents in order to efficiently achieved high quality herbal products. IL has a wide range of applications and has the ability to separate a wide range of organic and inorganic compounds. Other than their unique properties, it can be designed according to the specific needs by matching the anion and cation while considering their bonds and other properties exhibited. In addition, IL has been experimentally proven to extract higher amount of phytochemicals compared to conventional solvents [8].

This study is to select the highest potential IL solvent to extract main flavonoids and phenolic acids from *L. pumila* through a property prediction modeling approach. Main phytochemicals considered are kaempferol, myricetin and quercetin for the range of flavonoids and gallic acid and caffeic acid of phenolic acids range. These main phytochemicals were selected based on previous studies of [9] and [10].

As an extraction solvent, ILs are responsible to stimulate the interactions between the analyte and the solvent [11]. The types of anion and cations in the ILs determined the solubility of the targeted phytochemicals. Phytochemical extraction solvents involved in this study are various IL with imidazolium-based cation for alkyl chains ranging from ethyl to octyl and anions consisting of bis(trifluoromethylsulfonyl)imide [Tf$_2$N$^{-}$], tetrafluoroborate [BF$_4$]$^{-}$, hexafluorophosphate [PF$_6$]$^{-}$, bromide [Br$^{-}$], chloride [Cl$^{-}$], acetate [Ac$^{-}$], ethylsulfate [EtSO$_4$]$^{-}$, trifluoroacetate [CF$_3$COO]$^{-}$ and trifluoromethanesulfonate [CF$_3$SO$_3$]$^{-}$.

Due to thousands possibilities of IL anions and cations interactions, property prediction methods such as activity coefficient models, group contribution methods, COSMO-RS and molecular simulations to reduce the cost of experimentation [12]. However, activity coefficient models, group contribution methods and COSMO-RS still requires some experimental results to optimize their parameters and validate the findings.

The solubility parameter is one of the effective screening tools for solvent selection due to the basic principle ‘like dissolves like’ [13]. According to the principle, when the solubility parameter of the solvent and solutes approach with each other, the solubility of the solute is higher in that solvent. The solubility parameter can be estimated experimentally or by using theoretical methods [14][15]. The experimental methods include measuring the heat of vaporization by calorimetry, inverse gas chromatography (IGC), measuring melting temperatures, intrinsic viscosity measurements, or by using the activation energy of viscosity [16][17]. While, theoretical approach in determining solubility parameters include using intrinsic viscosity, enthalpy of vaporization, PC-SAFT such as COSMO [18]. However, some of these theoretical methods have some limitations and still requires parameter from experimental data.
2. Methodology

2.1. Stage 1: Database and Property Models Library Development

An Excel-based database has been developed, containing physicochemical data of IL and phytochemicals obtained and collected from literatures, ILThermo NIST database [19] and an unpublished IL database from Technical University of Denmark (DTU). The collected data include molecular weight, melting point, heat of fusion, viscosity, density and Hansen solubility parameters (HSPs).

Besides physicochemical data, property models with its parameters are also included in the database. An extensive review of applicable property models were done through literature search. Available models were collected and analysed further to determine the accuracy of the equations for IL/phytochemical systems. In the current state of the database, HSPs models for IL involving an anion-cation correlation model [15] and Marrero-Gani (MG) group contribution model [20] for phytochemicals have been included. For solid liquid equilibrium (SLE) calculations, parameters of the MG model for melting point and heat of fusion are also included with the UNIFAC-IL activity coefficient model [21]. For these models, missing parameters are predicted based on available data.

2.2. Stage 2: Solubility Parameter Determination

In this section, HSPs of the phytochemicals (kaempferol, myricetin, quercetin, gallic acid and caffeic acid) are determined using the MG model. Meanwhile, the HSPs of imidazolium-based ILs were estimated using the anion-cation correlation model. The MG model can be used to estimate the properties such as melting point, heat of fusion, HSP and boiling point of pure organic compounds such as phytochemicals. The mathematical expression is shown in equation (1) where $f(X)$ is a function of property $X$ (HSP in this study) and contain the additional adjustable model parameters and constants depending on the property involved.

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k E_k O_k$$

($C_i$ is the contribution of first order group that occurs $N_i$ times, $M_j$ is the contribution of second order group that occurs $D_j$ times and $E_k$ is the contribution of third order group that occurs $O_k$ times. Equation equation (2), (3) and (4) describe the partial solubility parameters based on MG group contribution model. The value of partial solubility of dispersion, $(\delta_d)$, polar $(\delta_p)$, and hydrogen bond $(\delta_h)$ were obtained using the Integrated Computer Aided System (ICAS) Version 18.0 software.

$$\delta_d = \left( \sum_i N_i \delta_{d1i} + \sum_j M_j \delta_{d2j} + \sum_k O_k \delta_{d3k} \right) MPa^{1/2}$$

$$\delta_p = \left( \sum_i N_i \delta_{p1i} + \sum_j M_j \delta_{p2j} + \sum_k O_k \delta_{p3k} \right) MPa^{1/2}$$

$$\delta_h = \left( \sum_i N_i \delta_{h1i} + \sum_j M_j \delta_{h2j} + \sum_k O_k \delta_{h3k} \right) MPa^{1/2}$$

where $\delta_{d1i}$, $\delta_{d2j}$, $\delta_{d3k}$, $\delta_{p1i}$, $\delta_{p2j}$, $\delta_{p3k}$ and $\delta_{h1i}$, $\delta_{h2j}$, $\delta_{h3k}$ denotes the group contribution parameter for first-order group, second-order group and third-order group respectively.

In this study, HSPs of ILs were determined by using a correlation of the weighted sum of attractions and repulsions between the anions and cations [15]. By considering IL as a mixture of cation and anion, using the concept of group contribution, the solubility parameters of various anions and cations have been computed through molecular simulations. The correlation model is shown in equation (5), where $\delta_n$ represents the HSPs of ILs, $\delta_{an}$ represents solubility parameter contribution for cations and $\delta_{an}$ represents solubility parameter of anions.
According to this correlation, HSPs of IL can be determined by having solubility parameter contributions of both cations and anions; or by the experimental solubility parameter of ionic liquids. Thus, this correlation can be used to determine wide range of ionic liquid with available data.

However, certain parameter such as for cations ([DMIM], [PMIM], [PeMIM] and [HeMIM]) were not available. Therefore, in this study, we determined the missing parameters using the systematic correlation/relationship between the molecular weight of cation and solubility parameters. The missing anions (acetate [Ac], ethylsulfate [EtSO₄], and trifluoromethanesulfonate [CF₃SO₃]) solubility parameter contribution were determined using anion-cation relationship using the experimental solubility parameter data.

Using both HSP values of the ILs and phytochemicals, the total solubility of phytochemicals in the ILs can be predicted. ILs candidates screened by the total phytochemicals solubility value. The model for total solubility of phytochemical in a solvent is shown as in equation (6).

\[
\Delta \delta = \left[ (\delta_{d,P} - \delta_{d,S})^2 + (\delta_{p,P} - \delta_{p,S})^2 + (\delta_{h,P} - \delta_{h,S})^2 \right]
\]

where \(\delta_{d,P}\) (MPa\(^{1/2}\)) is partial dispersion solubility of phytochemical, \(\delta_{d,S}\) (MPa\(^{1/2}\)) is partial dispersion solubility of solvent, \(\delta_{p,P}\) (MPa\(^{1/2}\)) is partial polar solubility of phytochemical, \(\delta_{p,S}\) (MPa\(^{1/2}\)) is partial polar solubility of solvent, \(\delta_{h,P}\) (MPa\(^{1/2}\)) is partial hydrogen bond solubility of phytochemical and \(\delta_{h,S}\) (MPa\(^{1/2}\)) is partial hydrogen bond solubility of solvent. According to the equation above, it can be concluded that small value of \(\Delta \delta\) indicates that the solvent is more efficient for the extraction of that compound [22].

Only ILs with total phytochemicals value below 10 Mpa\(^{1/2}\) which indicates 90% solubility of solute in a solvent were selected for further analysis. This value was selected according to published experimental data where phytochemicals recoveries were around 97 percent in ionic liquid solvent [23].

2.3 Stage 3: UNIFAC-IL Binary Interaction Parameter Regression and SLE Calculations

The SLE thermodynamic model was used together with MG model (to predict melting point and heat of fusion when the experimental data is missing) and UNIFAC-IL (to predict activity coefficient). The SLE of a system can be determined by involving pure component properties of the system such as melting point and enthalpy of fusion of the solid (solute) into the system’s real behaviour in the liquid phase [24] such as in equation (7).

\[
\ln(x_i \gamma_i) = \frac{\Delta h_{fus} (T_m)}{R} \left( \frac{1}{T_m} - \frac{1}{T} \right)
\]

where, \(x_i\) is mole fraction of dissolved solute, \(\gamma_i\) is activity coefficient of solute in solute-solvent system, \(\Delta h_{fus}\) is an enthalpy of fusion, \(R\) is a gas constant and \(T_m\) is the melting point.

The UNIFAC-IL model equation (8) consists of the combinatorial contribution (\(\gamma'_i\)), due to differences in molecular size and shape; and the residual contribution (\(\gamma''_i\)) arising due to the differences in intermolecular forces of attraction. The activity coefficient is expressed as functions of composition and temperature.

\[
\ln \gamma_i = \ln \gamma'_i + \ln \gamma''_i
\]
Experimental data, parameter regression work was conducted in order to extend the UNIFAC-IL parameter matrix. The parameter regression were done using the Thermodynamic Model Parameter Equation (the TML) module in the ICAS 18.0 software using available published experimental data. Regressed parameters were then used for the estimation of solubility of the selected phytochemicals in the screened ILs.

2.4 Selection of the Best Ionic Liquid Solvent Candidates
UNIFAC-IL was used together with the SLE equation to predict the solubility of flavonoids and phenolic acids solubility in the selected imidazolium based ILs. In Stage 2, the HSPs calculation had assisted in the screening and reducing the large number of ILs. Based on the screened ILs SLE was used to determine that amount of the flavonoids and phenolic acids that can be extracted by the imidazolium-based ILs. Based on the predicted solubility values of these phytochemicals, the ILs were ranked to determine the best solvent.

3. Results and Discussion
This study focused on determining the best imidazolium-based IL solvent for the extraction of phytochemicals (flavonoid and phenolic acid) from L. pumila. First it has been observed that the selections of 163 imidazolium-based IL are able to extract the main flavonoid and phenolic acid compound from the selected herbs [23][26]. In this case, the main flavonoid and phenolic acid identified from L. pumila are kaempferol, quercetin and myricetin for flavonoid; and gallic acid and caffeic acid of the phenolic acids group.

3.1. Prediction of Phytochemicals/ILs HSPs and Determination of Missing Cation and Anion Solubility Parameters.
Solubility parameter and other physicochemical properties for L. pumila phytochemicals are predicted by MG model and presented in Table 1. Phytochemicals properties such as solubility parameter, heat of fusion and melting point were determined using ProPred of the Integrated Computer Aided System (ICAS) Version 18.0 software.

According to previous study, increase of alkyl chain length in ILs enhanced the solubility efficiency of both flavonoid and phenolic acid in the extraction solvent [23]. This is due to the increasing molecular weight of the IL as extraction solvent [27]. On the other hand, solubility parameter of IL observed to be decreasing with the increasing molecular weight of the IL’s cation (Figure 1) and viscosity of the solvent candidates. The solubility parameter value decreased when the alkyl chain length and viscosity increased. The relationship can be included as HSPs equation as presented by equation (9).

\[
\delta = \left[ \frac{\Delta H_v - RT}{V_m} \right]^{0.5}
\]

where \(\delta\) is the solubility parameter, \(\Delta H_v\) is the heat of vaporization, \(R\) is a gas constant, \(T\) for temperature and \(V_m\) is the molar volume.
Figure 1. Relationship of IL solubility parameter with cation molecular weight

Table 1. Phytochemicals properties predicted by Marrero – Gani group contribution method.

| Phytochemicals | FLAVONOID | PHENOLIC ACID |
|----------------|------------|---------------|
|                | Kaempferol | Quercetin     | Myricetin     | Gallic Acid | Caffeic acid |
| Structure      | ![Structure](image) | ![Structure](image) | ![Structure](image) | ![Structure](image) | ![Structure](image) |
| Molecular weight (g/mol) | 286.24 | 302.24 | 318.23 | 170.12 | 180.16 |
| Melting point, T_m (K) | 503.59 | 520.29 | 537.53 | 492.89 | 443.75 |
| Boiling point, T_b (K) | 729.78 | 751.01 | 771.02 | 641.52 | 634.29 |
| Enthalpy of fusion (KJ/mol) | 120.69 | 133.7 | 146.71 | 89.66 | 81.05 |
| Hansen solubility parameter, δ (Mpa^1/2) | 32.07 | 34.29 | 36.51 | 28.82 | 28.34 |

However, no systematic relationship was found between anion molecular weight towards solubility parameter of IL. Figure 2 was constructed to show the anion’s influence on the solubility parameter of ILs based on 1, 3-dimethylimidazolium [DMIM] cations. The solubility parameter increases in the following order: [Tf_2N] < [CF_3SO_3] < [BF_4] < Br < [Cl] < [EtSO_4] < [Ac] < [CF_2COO] < [PF_6]. The highest values of δ^2 are for [EtSO_4], [Ac], [CF_2COO] and [PF_6] anions, whilst the lowest value is for the [Tf_2N] anion.
In the list of 163 imidazolium-based ILs, from the developed Excel database, it was found that 4 different cations of 1,3-dimethylimidazolium [DMIM], 1-methyl-3-propylimidazolium [PMIM], 1-methyl-3-pentylimidazolium [PeMIM] and 1-heptyl-3-methylimidazolium [HeMIM] with different alkyl chain length had an unknown solubility parameters, when checked with the available cation solubility parameters in the Sistla anion-cation correlation model. As observed, IL solubility parameters decreased linearly with the increasing value of the molecular weight of cation due to the increasing of alkyl chain as well. Based on the systematic relationship of the cation’s solubility parameter with its molecular weight as shown in figure 3, a fitted linear equation is obtained. Missing cation solubility parameters were determined from the equation based on its known molecular weight.

For the missing anion parameters, since no direct relationship can be found between its molecular weight and the IL solubility parameter, it is determined based on the relationship between cation solubility parameter and the solubility parameter of ILs containing the same cation and the missing anions (obtained from database). The relationship is based on equation (5). However, if the solubility parameter of the ILs containing related anion is not available, then the missing anion parameter cannot be determined.
Three new anions ([EtSO$_4$], [Ac] and [CF$_3$SO$_3$]) solubility parameter were determined. Table 2 recorded the solubility parameters of cations and anions (reported in [15] and determined from this study) used for phytochemical extraction of _L. pumila_ in this study.

**Table 2. Solubility parameters of cations and anions**

| Cation                        | δ cation | Anion       | δ anion |
|-------------------------------|----------|-------------|---------|
| *1,3-dimethylimidazolium      | DMIM     | [Tf$_2$N]   | 55.80   |
| 1-ethyl-3-methylimidazolium   | EMIM     | [BF$_4$]    | 73.25   |
| *1-methyl-3-propylimidazolium | PMIM     | [Br]        | 76.09   |
| 1-butyl-3-methylimidazolium   | BMIM     | [Cl]        | 76.44   |
| *1-methyl-3-pentylimidazolium | PeMIM    | [*[EtSO$_4$]] | 78.15   |
| 1-hexyl-3-methylimidazolium   | HMIM     | [*[Ac]]     | 78.32   |
| *1-heptyl-3-methylimidazolium | HeMIM    | [CF$_3$COO] | 84.70   |
| 1-methyl-3-octylimidazolium   | OMIM     | [PF$_6$]    | 85.80   |

* Cation with new predicted parameter value from this study

**3.2. Total Solubility of Phytochemicals**

Total solubility parameter is determined to distinguish the interactions of solute and solvent, thus the efficiency of the solvent used in extracting phytochemicals can be also identified [28]. Total solubility of _L. pumila_ flavonoid and phenolic acid was essential for solvent screening. According to equation (6), it can also be deduced that total solubility of phytochemical in a solvent can be determine by the difference of ionic liquid and phytochemical solubility parameter as shown in Table 3.

The total solvent candidates of 163 imidazolium-based ILs initially available in the database were screened down to 72 candidates due to the unavailable data of solubility parameter for ILs and also their anions. At this stage, by considering the total solubility of phytochemical in the solvent, solvent candidates were screened down to 51 ILs for flavonoid extraction and 32 ILs for phenolic acid extraction. This study consider total solubility less than value of 10 Mpa$^{1/2}$ due to previous published experimental studies of phytochemical extraction using ILs which have higher extraction value compared to organic solvent [26]. In this case, we only considered 32 screened ILs as we are targeting for IL solvent that efficiently extract both flavonoid and phenolic acids. Therefore, the screened candidates were further analysed using SLE relationship to determine the total mole fraction of solute that is soluble in the selected solvents.

**3.3. UNIFAC-IL Binary Interaction Parameter Regression and Solubility Determination from SLE**

As mentioned in Section 2.3, UNIFAC-IL was used to predict the activity coefficient term in the SLE equation to predict solubility. In order to use UNIFAC, the BIPs need to be available. Therefore the 32 screened ILs from Stage 2, were decomposed into smaller groups and the groups are compared with the published UNIFAC-IL parameter table. Based on analysis, it is found that there were still few missing parameters compared to UNIFAC parameter matrix of [25] such as [MIM][Tf$_2$N], [MIM][Ac], [MIM][Br] and [MIM][CF$_3$COO]. Therefore, these missing groups need to be added in the initial parameter table and their parameters need to be regressed from available experimental data.

$$F = K^r_i K^r_j \sum_i K^r_i \left( \frac{y_i - y_i^{Exp}}{y_i^{Exp}} \right)^2$$

(10)

The group BIPs were obtained by correlating the activity coefficients of binary systems involving
organic compounds in imidazolium based ILs. ICAS 18.0 was used for the BIPs regression with the objective function shown in figure 10, where $\gamma$ is activity coefficient of solute, $\gamma^\text{Exp}$ is the activity coefficient of solute obtained from the experimental data and $K_i$ is the equilibrium ratio (K-value).

Table 3. Total solubility parameter of phytochemicals in ionic liquids

| Ionic liquid | $\delta$ predicted (Mpa$^{1/2}$) | Total solubility, $\Delta\delta$ flavonoid (Mpa$^{1/2}$) | Total solubility $\Delta\delta$ phenolic acid (Mpa$^{1/2}$) |
|--------------|----------------------------------|-------------------------------------------------|-------------------------------------------------|
|              | Kaempferol | Quercetin | Myricetin | Mean | Gallic acid | Caffeic acid | Mean |
| [OMIM][Tf$_2$N] | 25.02 | 3.05 | 2.92 | 3.33 | 3.10 | 6.65 | 4.85 | 5.75 |
| [HeMIM][Tf$_2$N] | 25.33 | 3.36 | 3.23 | 3.64 | 3.41 | 6.96 | 5.16 | 6.06 |
| [HMIM][Tf$_2$N] | 25.66 | 3.69 | 3.56 | 3.97 | 3.74 | 7.29 | 5.49 | 6.39 |
| [OMIM][CF$_3$SO$_3$] | 25.98 | 4.01 | 3.88 | 4.29 | 4.06 | 7.61 | 5.81 | 6.71 |
| [HeMIM][CF$_3$SO$_3$] | 26.36 | 4.39 | 4.26 | 4.67 | 4.44 | 7.99 | 6.19 | 7.09 |
| [PeMIM][Tf$_2$N] | 26.39 | 4.42 | 4.29 | 4.70 | 4.47 | 8.02 | 6.22 | 7.12 |
| [BMIM][Tf$_2$N] | 26.55 | 4.58 | 4.45 | 4.86 | 4.63 | 8.18 | 6.38 | 7.28 |
| [HMIM][CF$_3$SO$_3$] | 26.76 | 4.79 | 4.66 | 5.07 | 4.84 | 8.39 | 6.59 | 7.49 |
| [PMIM][Tf$_2$N] | 27.15 | 5.18 | 5.05 | 5.46 | 5.23 | 8.78 | 6.98 | 7.88 |
| [EMIM][Tf$_2$N] | 27.56 | 5.59 | 5.46 | 5.87 | 5.64 | 9.19 | 7.39 | 8.29 |
| [OMIM][BF$_4$] | 27.62 | 5.65 | 5.52 | 5.93 | 5.70 | 9.25 | 7.45 | 8.35 |
| [DMIM][Tf$_2$N] | 27.64 | 5.67 | 5.54 | 5.95 | 5.72 | 9.27 | 7.47 | 8.37 |
| [PeMIM][CF$_3$SO$_3$] | 27.68 | 5.71 | 5.58 | 5.99 | 5.76 | 9.31 | 7.51 | 8.41 |
| [OMIM][Br] | 27.82 | 5.85 | 5.72 | 6.13 | 5.90 | 9.45 | 7.65 | 8.55 |
| [OMIM][Cl] | 27.84 | 5.87 | 5.74 | 6.15 | 5.92 | 9.47 | 7.67 | 8.57 |
| [BMIM][CF$_3$SO$_3$] | 27.88 | 5.91 | 5.78 | 6.19 | 5.96 | 9.51 | 7.71 | 8.61 |
| [OMIM][Et$_2$SO$_4$] | 27.92 | 5.95 | 5.82 | 6.23 | 6.00 | 9.55 | 7.75 | 8.65 |
| [OMIM][Ac] | 27.93 | 5.96 | 5.83 | 6.24 | 6.01 | 9.56 | 7.76 | 8.66 |
| [OMIM][PF$_3$] | 28.04 | 6.07 | 5.94 | 6.35 | 6.12 | 9.67 | 7.87 | 8.77 |
| [OMIM][CF$_3$COO] | 28.05 | 6.08 | 5.95 | 6.36 | 6.13 | 9.68 | 7.88 | 8.78 |
| [HeMIM][BF$_4$] | 28.17 | 6.20 | 6.07 | 6.48 | 6.25 | 9.80 | 8.00 | 8.90 |
| [HeMIM][Br] | 28.41 | 6.44 | 6.31 | 6.72 | 6.49 | 10.04 | 8.24 | 9.14 |
| [HeMIM][Cl] | 28.43 | 6.46 | 6.33 | 6.74 | 6.51 | 10.06 | 8.26 | 9.16 |
| [HeMIM][Et$_2$SO$_4$] | 28.55 | 6.58 | 6.45 | 6.86 | 6.63 | 10.18 | 8.38 | 9.28 |
| [PMIM][CF$_3$SO$_3$] | 28.70 | 6.73 | 6.60 | 7.01 | 6.78 | 10.33 | 8.53 | 9.43 |
| [HeMIM][CF$_3$COO] | 28.77 | 6.80 | 6.67 | 7.08 | 6.85 | 10.40 | 8.60 | 9.50 |
| [HMIM][BF$_4$] | 28.77 | 6.80 | 6.67 | 7.08 | 6.85 | 10.40 | 8.60 | 9.50 |
| [HeMIM][PF$_3$] | 28.77 | 6.80 | 6.67 | 7.08 | 6.85 | 10.40 | 8.60 | 9.50 |
| [HMIM][Br] | 29.05 | 7.08 | 6.95 | 7.36 | 7.13 | 10.68 | 8.88 | 9.78 |
| [HMIM][Cl] | 29.08 | 7.11 | 6.98 | 7.39 | 7.16 | 10.71 | 8.91 | 9.81 |
| [HMIM][Et$_2$SO$_4$] | 29.22 | 7.25 | 7.12 | 7.53 | 7.30 | 10.85 | 9.05 | 9.95 |
| [HMIM][Ac] | 29.22 | 7.25 | 7.12 | 7.53 | 7.30 | 10.85 | 9.05 | 9.95 |

Table 4. Volume, $R_k$ and Surface Area, $Q_k$ for parameter regression

| New subgroups | $R_k$ | $Q_k$ |
|---------------|-------|-------|
| [MIM][CF$_3$COO] | 7.22 | 1.67 |
| [MIM][Et$_2$SO$_4$] | 9.37 | 2.01 |
Table 4 described the value for volume, $R_k$ and surface area, $Q_k$ for the new sub-groups and table 5 showed the tabulated regressed $a_{mn}$ and $a_{nm}$ BIPs with tolerance of $1 \times 10^{-6}$ and maximum iteration steps of 1000. The table listed the value of missing parameters as an additional to the UNIFAC parameter matrix of [25]. However, parameter regression were not done for [MIM] [EtSO$_4$] due to the lack of experimental data point. As according to the BIPs, there were only interaction between organic group with ILs group and none interaction between the ILs group in the system. All values were in the range to fit the parameters.

The activity coefficient estimated using UNIFAC-IL using the available and newly regressed parameters were obtained and applied to the SLE equation (7). The solubility of flavonoid and phenolic acid in 5 different type of ILs (the most optimum from the list of 32 ILs) are presented in Figure 4. As observed, ILs with [HMIM] [Tf$_2$N], [OMIM][Tf$_2$N] and [HeMIM][Tf$_2$N] anions and longer alkyl chain were the best solvents for extracting flavonoid and [HMIM][CF$_3$SO$_3$] and [OMIM][CF$_3$SO$_3$] for phenolic acid. ILs with organic fluorinated anions tend to have higher electrostatic contribution (equal to the collective contribution of coulombic and hydrogen bonding) and higher thermal stability [15].

In addition, the increase of alkyl chain length increased the hydrophobicity of the ionic liquid [29], therefore suitable in flavonoid and phenolic acid extraction. Furthermore, according to [30], the solubility of the solutes increases as the length of alkyl substitute on the ionic liquid increases as this study showed that the longer alkyl chain had the highest value of solubility data.

Table 5. UNIFAC-IL group interaction parameters ($a_{mn} \neq a_{nm}$)

|        | CH2  | C=C  | ACH  | ACOH | OH   | COOH | [MIM] [Tf$_2$N] | [MIM] [Ac] | [MIM] [Br] | [MIM] [CF$_3$COO] |
|--------|------|------|------|------|------|------|----------------|------------|------------|-------------------|
| CH2    |      |      |      |      |      |      | 523.31         | 49.07      | 20.80      | 0.00              |
| C=C    |      |      |      |      |      |      | -0.17          | -13.78     | 23.17      | -1.87             |
| ACH    |      |      |      |      |      |      | -686.62        | -31.72     | -38.15     | 0.26              |
| OH     |      |      |      |      |      |      | 6.96           | -25.93     | -50.98     | 0.00              |
| ACOH   |      |      |      |      |      |      | -756.67        | -52.55     | -50.83     | -10.08            |
| COOH   |      |      |      |      |      |      | -2850.23       | -2.03      | 10.62      | 0.00              |

-220.02 | 0.27 | -2774.24 | -0.09 | -3243.87 | -2410.81 |
-61.51  | -0.01 | -32.23  | -21.62 | -94.66  | -6.21   |
71.58   | -11.21 | -0.78   | -25.94 | -101.91 | -7.43   |
0.02    | 0.04  | 21.45   | -2.20  | -12.86  | -7.91   |

- No interactions between group parameters
- According to previously published parameters from [25]
Figure 4. Solubility of a) Kaempferol b) Quercetin c) myricetin d) gallic acid and e) caffeic acid in six different ionic liquids

4. Conclusions
Anion-cation relationship is a convenient method to determine an unknown solubility parameter of IL, with available data of either anion and/or cation and/or solubility parameter of ionic liquid. In this study, four solubility parameters of cation ([DMIM], [PMIM], [HMIM] and [HeMIM]) and three solubility parameter of anion ([EtSO₄], [Ac] and [CF₃SO₃]) were added into the database using data regression. Solubility parameter is important for screening and reducing solvent candidates. Screened candidates were further assessed using UNIFAC-IL to determine the best IL candidates to extract flavonoid and phenolic acid from *L. pumila*.

5. References
[1] Doughari J H 2012 *IntechOpen* 1–33
[2] Zhang Q and Ye M 2008 J Chromatogr A 1216 1954–69
[3] Azmir J, Zaidul I S M, Rahman M M, Sharif K M, Mohamed A, Sahena F and Omar K M 2013 *Journal of Food Engineering* 117 426–36
Acknowledgement

The financial supports from Fundamental Research Grant Scheme (FRGS), Malaysia Ministry of Higher Education (Vote number: R.J130000.7844.4F670) and MyBrain15 scholarship by Ministry of Higher Education Malaysia is greatly appreciated. Data on IL supplied by Prof. Rafiqul Gani from Department of Chemical and Biochemical Engineering, Technical University of Denmark is highly appreciated.