

SYSTEMATIC METHODOLOGY FOR THE DESIGN OF BINARY SOLVENT BLENDS FOR EXTRACTION OF HERBAL PHYTOCHEMICALS WITH COST EVALUATION

Siti Nuurul Huda Mohammad Azmin^{a*}, Mohd Shukri Mat Nor^{b,c}

^aFaculty of Agro-Based Industry, Universiti Malaysia Kelantan Jeli Campus, Locked Bag 100, 17600 Jeli Kelantan, Malaysia.

^bProcess Systems Engineering Centre (PROSPECT), Research Institute of Sustainable Environment (RISE), Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.

^cFaculty of Chemical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.

* Corresponding author: E-mail: huda.ma@umk.edu.my
Tel.: +609-9477154

ABSTRACT

The trial-and-error solvent selection method to obtain herbal phytochemicals is time consuming and limited by effort and cost. The combination of property prediction models with computer-assisted search is one way to overcome these drawbacks. Thus, the main objective of this work is to present a computer-aided methodology for the design of solvent blends in extracting herb phytochemicals optimally with cost evaluation. The methodology can be summarised into two main stages, namely, model-based design and experimental-verification stages. The result discussed in this paper is only for the first stage. The extraction of kaempferol from Kacip Fatimah herb is used as a base case study that follows all of the listed tasks. Five optimal binary solvent blends have been identified namely, methanol:isobutyraldehyde (M:IB), methanol:n-propionaldehyde (M:PP), methanol:water (M:W), methanol:ethyl acetate (M:EA) and methanol:acetic acid (M:AA). The M:IB solvent blend is able to extract the highest kaempferol yield while M:PP gives the highest profit.

Keywords: Herbal extraction, Kacip Fatimah, phytochemicals, extraction, product design, solvent blend.

1.0 INTRODUCTION

Herbal plants have been used throughout human history as sources of food, beauty enhancers and fragrances (Kumoro and Hasan, 2008). These plants have often been used to cure a broad range of acute and chronic conditions (Diallo *et al.*, 1999), which suggests that a vast wealth of knowledge could be gained through exploration of these plants. Phytochemical constituents in herbs, such as flavonoids, phenolics and saponins, are believed to have diverse therapeutic abilities and are able to reduce the risk of multiple diseases, including inflammatory conditions and cancer (Karimi *et al.*, 2013). Aqueous extracts of the Malaysian herb *L. pumila*, commonly known as Kacip Fatimah, have been demonstrated significantly to protect human dermal fibroblasts from cell damage caused by ultraviolet (UV) irradiation (Choi *et al.*, 2010), most likely as a result of the presence of flavonoids (Norhaiza *et al.*, 2009). The reported benefits have

contributed to the high demand for herbal products. Thus, the identification of the most effective methods for extracting the phytochemicals from plants has become increasingly important (Abdullah *et al.*, 2012).

Simple preparation methods for processing the plant herbs by boiling the entire plant or selected parts of the plant in water have been traditionally used by herbal medical practitioners since ancient times. As the herb is soaked in the solvents, the phytochemicals will diffuse out from the herbal plant cell to the solvent medium. The boiling method can be used but it is time consuming as there is no other driving force except heat that will increase the diffusion of these phytochemicals. Currently, there are many methods that have been applied to obtain these valuable phytochemicals, such as microwave-assisted extraction (MAE), sonication extraction, supercritical fluid extraction and so on, which have more than one driving force. For example, the MAE method will use heat and microwave power as an additional driving force to increase the diffusion rate of phytochemicals. In addition, these driving forces also will increase the amount of extracted phytochemicals. All of these modern techniques use solvent as a transfer medium of phytochemicals. Every solvent will attract different phytochemicals as the “like dissolve like” theory is implemented (Barton, 1990). This is the reason why different solvents will give different results (yields) in the extraction process.

In the case of herbal extraction, the main issue that needs to be improved with the current solvent selection method for extraction is the trial-and-error approach. The drawbacks of this solvent selection method are effort, cost and time (Samudra and Sahinidis, 2013, Karunanithi *et al.*, 2005). For example, the trial-and-error method requires the use of a series of preliminary experimental studies that involve raw materials (solvent; dried and ground herbal plant) and energy consumption (heat and electricity). Then, the result from this trial-and-error method will be used to choose a suitable solvent before real experiments are conducted. Another disadvantage of the trial-and-error method is that the number of experimental repetitions is unknown and might require a lot of time. In addition, traditional methods have focused on experiments using classes of solvent (polarity) with different classes giving rise to different phytochemical extractions (Karunanithi *et al.*, 2009). To extract one type of phytochemical, at least six solvents may be needed (Kerton and Marriott, 2013). If the number of solvents can be reduced, the amount of waste can be minimized, productivity can be maintained or increased and extraction time can be reduced (Kerton and Marriott, 2013). The combination of property-predictive models with computer-assisted search is one way to overcome these drawbacks (Samudra and Sahinidis, 2013). Most of the chemical products are designed through experimentally based trial-and-error techniques (Gani, 2004). Hence, computer-aided techniques for chemical product design have been developed (Gani and Brignole, 1983, Garg and Achenie, 2001). Computer-aided technique is defined by Karunanithi *et al.* (2005) as “given a set of chemicals and specified set of property constraints, determine the optimal mixture”. A way to solve this problem is to employ a systematic methodology consisting of a sequence of work flows that guides the user to get the needed product. The main objective is to match the desired product properties within the set target limit for every selected property. The computer-aided techniques also can speed up the process of designing a product and this method is more efficient because the validated chemical product models can be applied (Gani, 2004). However, the main challenge in this

approach is the availability of chemical product models that could be applied with acceptable errors. In addition, a chemical product cannot be designed without considering the profitability of the product (Hill, 2009, Lee *et al.*, 2014).

The computer-aided approach has been applied mostly in designing solvent for separation of solid–liquid and liquid–liquid systems. Previously, Karunanithi *et al.* (2005) designed an optimal solvent for the separation of acetic acid from water using liquid–liquid extraction (LLE) based on a computer-aided molecular/mixture design (CAMD), while Folić *et al.* (2005) also used the same method for the design of solvents involved in reaction processes. The CAMD framework has also been applied to design solvents for crystallization of pharmaceutical compounds by Karunanithi *et al.* (2006), whereas Cheng and Wang (2010) used this method to find a feasible biocompatible solvent for an extractive fermentation. Conte *et al.* (2012) designed a solvent blend for the formulation of paint and insect repellent using model-based computer-aided tools. In addition, Cheng and Wang (2007) applied mixed-integer hybrid differential evolution to design an optimal biocompatible solvent for an extractive fermentation process. Mathematical modelling was used by Vanderveen *et al.* (2014) to design switchable-hydrophilicity solvents. The designed solvents can be used as solvent removal that does not require the use of volatile compounds in distillation systems. The mathematical modelling also was used to design a solvent in a coking wastewater treatment process by Liao *et al.* (2014), while Cheng and Wang (2008) used mathematical modelling to design a solvent for ethanol extractive fermentation with cell recycling. In addition, the same cases can also be approached using different methods. As an example, Damartzis *et al.* (2014) applied mathematical modelling, while Stavrou *et al.* (2014) used continuous molecular targeting–computer-aided molecular design (CoMT–CAMD) to design an optimal solvent based on the post-combustion of carbon dioxide capture process. Papadopoulos and Seferlis (2009) proposed a systematic approach for designing solvents in the separation of liquid–liquid mixtures using extractive distillation based on solvents' economic and behavioural characteristics. Table 1 summarizes the developed methods for various applications and case studies. In spite of using all of the previous methods, the validation of methodology used must be performed with selected case studies. This important step will determine whether the method is applicable or not with the current cases and the possibility of extending it to other cases.

This study proposes a framework in designing binary solvent blends to extract the targeted herbal phytochemicals. The proposed framework involves five levels. Level 1 is screening the pure component properties of solvents, Levels 2 and 3 consider linear and non-linear constraints, respectively, Level 4 is stability analysis and Level 5 is calculation of the cost and profit. Phytochemicals and solvent properties relationship will be evaluated as the aim is to design solvent blends that can extract the maximum amount of phytochemicals from herbs. The solvent to be designed will consider all of the safety, economic and environmental issues. The main objective of this work is to develop a new solvent blend for the maximum extraction of herbal phytochemicals using a computer-aided approach with cost evaluation. The method is demonstrated by designing a solvent blend for Kacip Fatimah herb as a base case study.

Table 1: Works on solvent design with their case studies

Method used	Case study	References
CAMD	Solvent for the separation of acetic acid from water	Karunanithi <i>et al.</i> (2005)
	Solvent involved in reaction processes	Folić <i>et al.</i> (2005)
	Solvents for crystallization of pharmaceutical compounds	Karunanithi <i>et al.</i> (2006)
	Feasible biocompatible solvent for an extractive fermentation	Cheng and Wang (2010)
Model-based computer-aided tools	Formulation of paint and insect repellent	Conte <i>et al.</i> (2012)
Mixed-integer hybrid differential evolution	Optimal biocompatible solvent for an extractive fermentation process	(Cheng and Wang, 2007)
Mathematical modelling	Switchable-hydrophilicity solvents that can be used as solvent removal	Vanderveen <i>et al.</i> (2014)
	Solvents in coking wastewater treatment process	Liao <i>et al.</i> (2014)
	Solvent for ethanol extractive fermentation with cell recycling	Cheng and Wang (2008)
	Solvent based on post-combustion of carbon dioxide capture process	Damartzis <i>et al.</i> (2014)
Continuous molecular targeting-computer-aided molecular design (CoMT-CAMD)	Solvent based on post-combustion of carbon dioxide capture process	Stavrou <i>et al.</i> (2014)
Systematic approach	Solvents in the separation of liquid-liquid mixtures using extractive distillation	Papadopoulos and Seferlis (2009)
	Fuel additives that are converted from biomass	

2.0 METHODOLOGY

This systematic methodology employs the reverse design approach (Gani, 2004), where the targets of the design problem are defined and the solvent blends that match the targets are identified. This reverse design approach has been chosen because it is ideally suited to handle “define target-match target” problems. In addition, the approach is able to manage the complexity of the design problem efficiently and to reduce the search space (Karunanithi *et al.*, 2005). Normally, a set of solvents is systematically generated and screened. Note that, in this solvent design methodology, only binary mixtures are considered, but it can be easily extended to multicomponent mixtures.

2.1 Systematic methodology of solvent design for phytochemical extraction

In this study, the overall methodology is divided into two stages, namely, the model-based design and experimental verification stages, as shown in Figure 1.

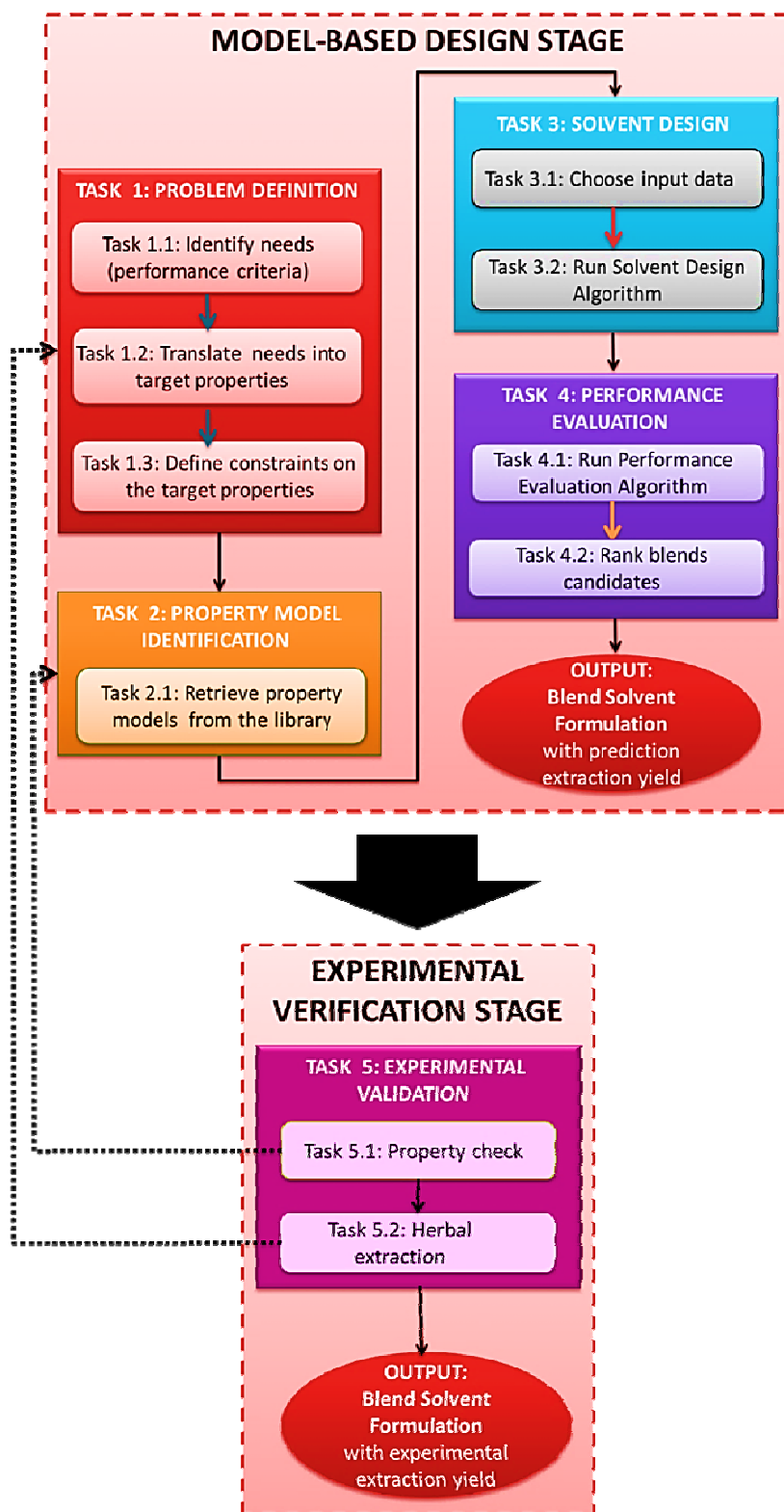


Figure 1: Systematic methodology for solvent design in extracting phytochemicals from herbs

In the model-based design stage, it is separated into four main tasks. Task 1 is problem definition, where this task would identify the user needs, translate the user needs into targeted properties and obtain the targeted property constraints. Task 2 is the property model identification, where the models to be used to estimate the targeted properties are

retrieved from the property models library (see Section 2.1.2). Task 3 is specifically for solvent design, whereas Tasks 1 and 2 are used as inputs in the solvent design algorithm. The objective of this task is to find solvent blend candidates that satisfy the solvent target properties at low cost. This algorithm must be tested with the case study to make sure that the algorithm is suitable with the current application. Task 4 is to evaluate the performance of the designed solvents. In this task, the optimal amounts of the phytochemicals that can be extracted by the designed solvents are estimated using an appropriate prediction model. Finally, the results obtained from the model-based design stage need to be verified experimentally at the experimental verification stage. The validation task is divided into two steps. The first step is to validate the selected solvent properties and it is then followed by the experiment on herbal extraction using the designed solvents. The second step is to validate the actual amount of the phytochemicals that can be extracted using the designated solvents. Both experimental results are then used to calculate the error between the predicted and experimental values. If the experimental validation for the first part gives a huge error, the task amendment must be done starting from Task 2 (model identification). Lack of success for the first part means that either a high error value is obtained (comparison between the prediction and experimental values for every selected property) or the property values are outside the set property constraints.

If the experimental validation for the second part also gives a huge error, the task amendment must be done starting from Task 1 (problem definition). The problem might originate because of unsuitable property selections, property constraints or property models with the current tested case study that makes the selection of solvent blends inappropriate.

2.1.1 Task 1: Problem definition

This task is divided into three sub-tasks, namely, identify needs, translate needs into target properties and define constraints of the target properties. The output of this task is a list of properties considered in the design with their target boundaries or constraints.

Task 1.1 Identify needs (performance criteria). This task is very important because it will determine the main function of the solvent to be designed. The needs are determined by defining the main functions of solvent in the herbal extraction process. For example, the solvent used in herbal extraction must be able to extract almost all phytochemicals from herbal plant or only one of the targeted phytochemicals. Thus, the needs here are to find what factor would influence the phytochemicals to diffuse from the herbal cell to the solvent medium. For this reason, the relationship of target phytochemicals and solvent used are very important. The most important criterion that must be emphasized is the diffusion mechanism of phytochemicals from the herb cell to the solvent medium, which can enhance the phytochemicals' extraction yields. The main factor that affects the diffusion mechanism is the solubility of phytochemicals in the selected solvent. This solubility is measured using the solubility parameter property. The solubility of phytochemicals increases as the solubility parameters of the phytochemicals and solvent used become closer to each other. This methodology, however, may not consider all of the factors because not all factors involved can be measured by properties. For example, even though the size of the raw material (herbal plant) would influence the extraction yield, it could not be measured by any property. Thus, the raw material size could not be considered in this methodology. Therefore, this

methodology only considers all the factors that can be measured by properties. The solvents used for herb phytochemicals extraction also must have specific criteria, such as they must be miscible with each other in order to prevent the formation of two layers of solvents, which will cause the extraction to become inefficient. In addition, the solvent must also be easily separated from the phytochemicals to produce high purity phytochemicals and enable the solvent to be reused for the same purpose.

Task 1.2 Translate needs into target properties. In this task, all of the needs that have been identified in Task 1.1 must be translated into target properties in order to evaluate them. This task requires knowledge for translation of needs into target properties. This knowledge is gathered and collected from the experts in the particular product or process as well as from patents and the literature (Lee *et al.*, 2014). They are put into a database called the knowledge base. The needed knowledge for this task is shown in Table 2.

Table 2: Needs, performance criteria and target properties employed in this study

Needs	Performance Criteria	Target Properties
Solvents must have almost the same value of polar property in order to make sure that phytochemicals can diffuse out from the herbal cell to the solvent medium.	Polarity	$\text{Log } p$
Solvent that would be designed must not evaporate easily to make sure that in the extraction process, the solvent does not dry.	Boiling point	T_b
Solvent that would be designed must not be too difficult to evaporate as it will be removed after the extraction process to get the crude extract.		
Safety of the solvent to the users/ researchers must be considered.	Toxicity	LC_{50}
Compatibility of the solvent to the phytochemicals must be considered so that the solvent will be easy to handle and effectively extract the target phytochemicals.	Solubility	δ
The solvent components must be miscible with each other to make sure that there is no undesired layer (immiscibility of solvent) in the extraction.	Miscibility/Stability	ΔG^{mix}
Solvent blend must not be too sticky as it will be removed from the mixture of solvent and herbal crude extract.	Viscosity	μ
The solvent that will be designed must not only be considered in term of effectiveness but also economics.	Price	C

Table 2: (continued)

Needs	Performance Criteria	Target Properties
As existing solvent is sold by price per weight, so the density of the blend solvent is also needed.	Density	ρ

Note that not all product needs can be translated into physicochemical properties. As an example, even though the herbs' planting location would influence the extraction yield, it could not be measured by any property. The knowledge base about the principle of herbal extraction must be used in this task. Then, all of the needs are translated into the target properties. These target properties could be classified into three terms, namely, properties that determine: 1) the main function of solvent (properties related to the choice of phytochemicals); 2) the solvent performance (physicochemical properties of the solvent); and 3) the solvent phase stability. Table 2 summarizes the needs, performance criteria and target properties employed in this study.

Task 1.3 Define constraints on the target properties. After the target properties in Task 1.2 are obtained, the constraints for all target properties must be specified. The constraint values can be obtained from the existing solvents used in the extraction process. These values are used as a benchmark in designing a base case of solvent. They could be changed to allow improvement for a better solvent design. Some of the constraint values are obtained from legislation or local regulations. For example, the vapour pressure of solvent blends must be less than a certain limit set by legislation to ensure that the solvent does not easily release into the environment (Klein *et al.*, 1992). All of the gathered target constraints are stored in the knowledge base. This limit is important in order to make sure that all designed blend solvents are within this set range.

2.1.2 Task 2: Property model identification

In this study, the physicochemical properties of interest are mixture properties because the solvent to be designed is in mixtures/blends. This requires mixture property models to estimate the mixture properties. The property models to estimate pure properties may also be needed if the pure properties are not available. Note that, whenever the experimental data are available, these data are directly employed in the calculations. The models to estimate both pure and mixture properties were collected from the literature and stored in a property models library, as listed in Tables 3 and 4. This library is created to store the related models that might be applicable for designing solvent blends. For pure component properties, it has been classified into primary and secondary properties. Primary properties could be directly determined using the group contribution (GC) method and only depend on molecular structure. Secondary properties could not be determined directly using the GC method, but it uses some of the primary properties as specified variables. As an example, the density calculation is a function of critical pressure, P_c , and temperature, T_c , in which both P_c and T_c could be determined using the GC method.

Task 2.1 Retrieve models from the library. In this task, the needed property models to solve the solvent design problem are retrieved from the library. These property models

would be used to predict the target property values where it is easier and faster than performing experiments on the solvents.

2.1.3 Task 3: Design solvent blend

Task 3 focuses on searching the solvent candidates that satisfy all of the property constraints set in Task 1.3. The objective of this task is to generate and screen all solvent blend candidates. In this task, the main part is the solvent design algorithm (Task 3.2), where the solvent blend candidates are generated and screened through five levels. Before performing the algorithm, the input data must be chosen. This algorithm gives a result as a list of all potential solvent blend candidates that could be used in herbal extraction with the cost of 1 ml of solvent blend. Then, all of the feasible solvent blends are ranked in ascending order based on solvent cost obtained in Task 3.2. This task was performed using MATLAB software.

Task 3.1 Choose input data. Results from Task 1 (Task 1.1: identify needs, Task 1.2: translate needs into target properties and Task 1.3: define constraints on the target properties) and Task 2 (Task 2.1: retrieve property models from the library) are used as inputs for this algorithm. In addition, selected phytochemical properties, the list of solvents with their associated properties and temperature that would be considered in the extraction process are also used as input data. Two types of input data must be listed, namely, solvent and phytochemical data. Both data must have all pure component property values for each solvent and phytochemical that affect the extraction of herb phytochemicals.

Task 3.2 Run solvent design algorithm. The solvent blend design algorithm employs a decomposition-based solution strategy where the number of feasible mixtures is systematically decreased in subsequent levels (Karunanithi *et al.*, 2005, Conte *et al.*, 2011, Yunus *et al.*, 2014). Figure 2 shows the solvent design algorithm that has been applied in this study. The input information for the algorithm are solvent and phytochemical data (necessary pure compound properties are stored), both linear and non-linear property models, constraints on target properties and designed temperature. As shown in Figure 2, the algorithm has five levels and they are ranked according to the hierarchy of calculations with increasing complexity. Level 1 is for screening pure component target properties, while Levels 2 and 3 solve the linear and non-linear target property constraints, respectively. Level 4 is to analyse the stability of mixture solvents and Level 5 is to calculate the cost of use for all potential solvent candidates. Results for every level would be continued to the next level, where the result for the previous level would be the input data for the next level. As an example, the result for Level 1 is a binary mixture that satisfies Step 1-1 (S1.1). Then this result would be the input data for Level 2.

The results from this algorithm are the solvent blends formulation with their composition, target property values and solvent blends cost. This algorithm is described and highlighted for binary solvent mixtures, but it can be extended to multicomponent solvent mixtures. A binary solvent mixture is a combination of two solvents, *i* and *j*, from solvent input data. Subscripts *i* and *j* represent the number of each solvent, where *i* is always less than *j* to avoid any repetition of formulation in binary solvent mixtures. As an example, input data contains three solvent candidates namely solvent 1, solvent 2 and solvent 3. If the binary solvents are to be designed, the possible combinations of

solvent are solvent 1 with solvent 2, solvent 1 with solvent 3 and solvent 2 with solvent 3. There will be no repetition of the same solvent combination in the reverse direction, such as solvent 2 with solvent 1, solvent 3 with solvent 1 and solvent 3 with solvent 2.

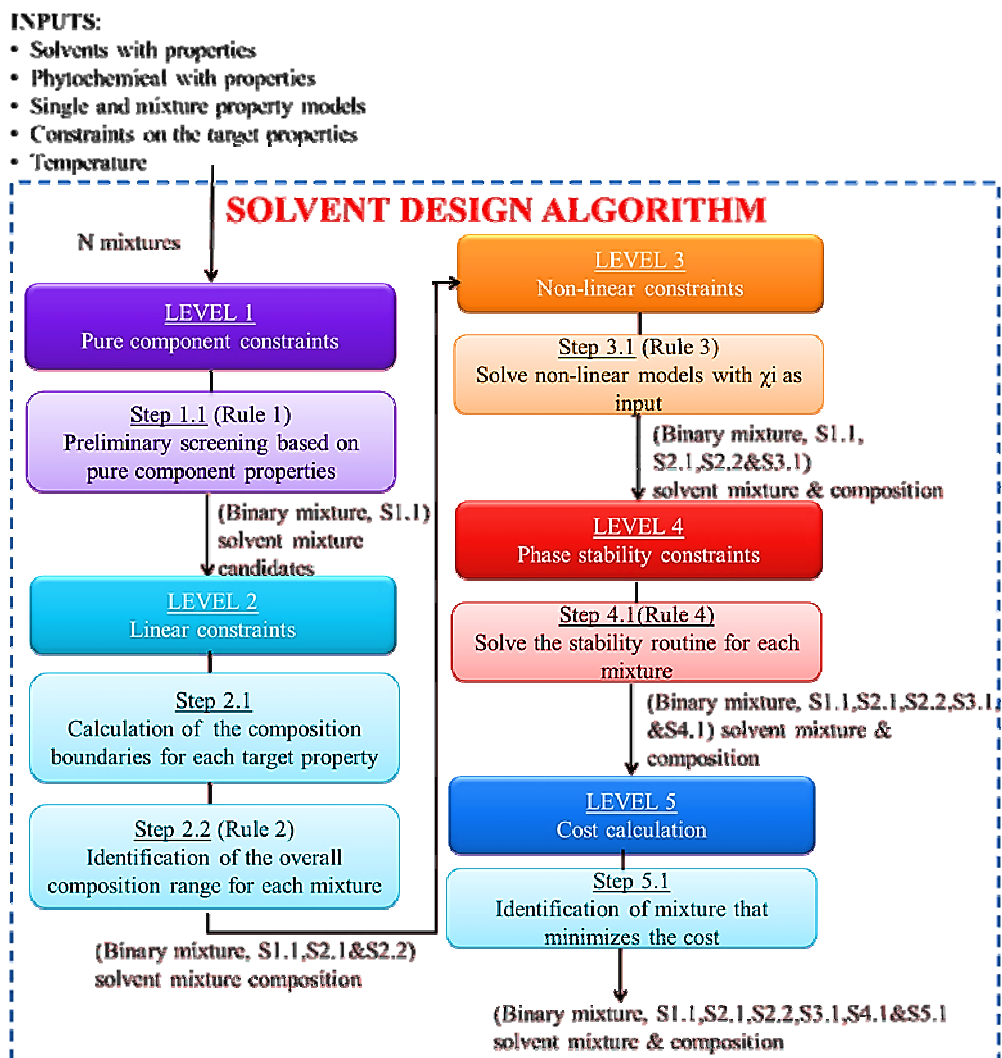


Figure 2: Solvent design algorithm

(i) Level 1: Pure component constraints. Level 1 will only screen on the pure component constraints. Pure component here means all of the considered phytochemical properties that could be related to the solvent properties.

Step 1.1: Preliminary screening based on pure component properties. In this step, target properties, values for ζ^k for each solvent component in the mixture, are compared with their set target values boundaries, ζ_{LB} and ζ_{UB} for each property, k . Figure 3 illustrates a comparison of solvent property values in the binary mixture, where ζ_{S1}^k represents the target property of solvent 1, ζ_{S2}^k is the target property of solvent 2, ζ_{LB}^k is the lower bound and ζ_{UB}^k is the upper bound of the target property, k . The target region is determined by using the target phytochemical property value. If one of the solvents in a mixture is in the target region, the mixture would be continued to the next level. Figure 3 shows the condition in which the solvent mixture satisfies the set condition.

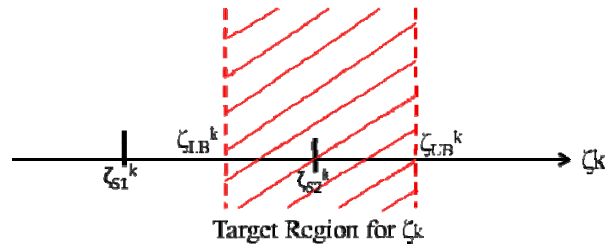


Figure 3: Pure property boundary

Rule 1: If the property values of every solvent in a mixture are both either lower than the lower bound values ($\zeta_{S1}^k < \zeta_{LB}^k$ and $\zeta_{S2}^k < \zeta_{LB}^k$) or greater than the upper bound values ($\zeta_{S1}^k > \zeta_{UB}^k$ and $\zeta_{S2}^k > \zeta_{UB}^k$), that solvent mixture would be rejected.

(i) Level 2: Linear constraints. Binary mixture screening of solvents starts at this level. Linear constraints are related to the properties described by a linear model. In this case study, linear model means following the linear mixing rule to compute the mixture target property. For a binary mixture, the generic form of the linear model is:

$$\zeta^{k,m} = \sum_{i=1}^{NS} x_i \zeta_i^{k,m} = x_1 \cdot \zeta_1^{k,m} + (1 - x_1) \cdot \zeta_2^{k,m} \quad (1)$$

In this equation, subscripts 1 and 2 indicate solvents 1 and 2 in the binary mixture, $\zeta_i^{k,m}$ is the pure solvent property k of compound i in the mixture m , x_i is the mole fraction of compound i , NS is number of solvents in the mixture.

Step 2.1: Calculation of the composition boundaries for each target property. In this step, the composition boundaries for each target property of solvent in the binary mixture are calculated using Equation (2). $\zeta^{k,m}$ is a specific target value for property k .

$$x_1^{k,m} = \frac{\zeta^{k,m} - \zeta_{S2}^{k,m}}{\zeta_{S1}^{k,m} - \zeta_{S2}^{k,m}} \quad (2)$$

As illustrated in Figure 4, the pure properties of solvents 1 and 2, $\zeta_{S1}^{k,m}$ and $\zeta_{S2}^{k,m}$ are in the target region, as shown in the shaded area. The dashed lines represent the target region limited by the upper, ζ_{UB}^k , and lower, ζ_{LB}^k , bounds. In this figure, the pure property of solvent 1 (S1) is less than that pure property of solvent 2 (S2). Thus, the property k for the solvent mixture must lie in the target region to match the design target.

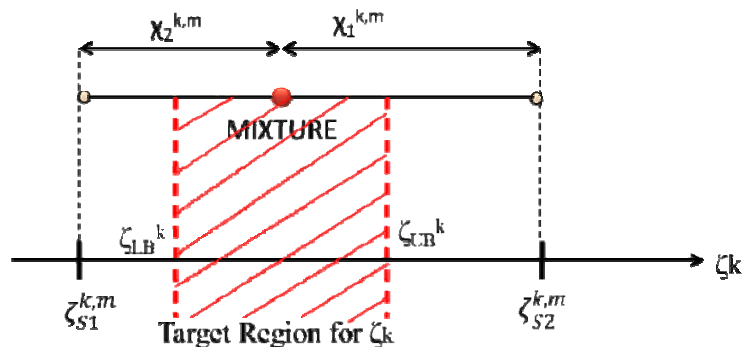


Figure 4: Mixture solvent boundary

Then, the composition range of solvent 1, ($x_{1,UB}^{k,m}$ and $x_{1,LB}^{k,m}$) for a binary mixture, m is calculated using Equations (3) and (4):

$$x_{1,UB}^{k,m} = \frac{\zeta_{LB}^k - \zeta_{S2}^k}{\zeta_{S1}^k - \zeta_{S2}^k} \quad (3)$$

$$x_{1,LB}^{k,m} = \frac{\zeta_{UB}^k - \zeta_{S2}^k}{\zeta_{S1}^k - \zeta_{S2}^k} \quad (4)$$

Step 2.2: Identification of the overall composition range for each mixture. Identification of the overall composition range ($x_{1,UB}^{k,m}$ and $x_{1,LB}^{k,m}$) for each mixture could be achieved by comparing the composition range of all target properties. As an example, if two different composition ranges for properties 1 and 2 are analysed, the overall composition range is the region where both composition ranges overlap. An illustration of this example is shown in Figure 5.

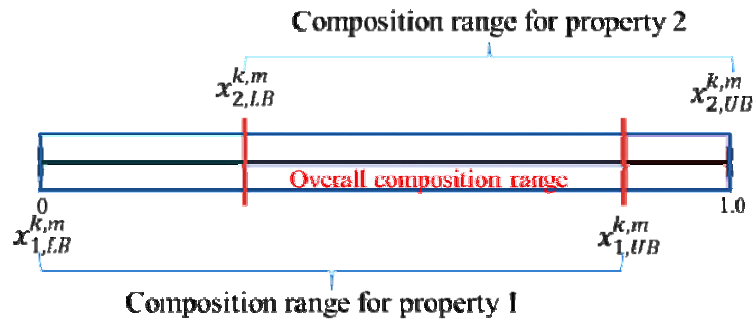


Figure 5: Overall composition range for a solvent mixture

As shown in this figure, the lower bound composition, $x_{1,LB}^{k,m}$, will be the lower bound of the composition range for property 2, while the upper bound composition, $x_{1,UB}^{k,m}$, will be the upper bound of the composition range for property 1. The minimum and maximum values of $x_{1,LB}^{k,m}$ and $x_{1,UB}^{k,m}$ for each property k are calculated using Equations (5) and (6):

$$x_{1,LB}^m = \max(x_{1,LB}^{k,m}) \quad (5)$$

$$x_{1,UB}^m = \min(x_{1,UB}^{k,m}) \quad (6)$$

Mathematically, the overall composition calculations follow Equations (3) and (4). The minimum and maximum value of $x_{1,UB}^{k,m}$ and $x_{1,UB}^{k,m}$ are calculated using Equations (5) and (6).

Rule 2: The solvent mixtures with the composition range of each property that do not overlap each other are rejected. Mathematically, any binary mixture that satisfies this condition, $x_{1,UB}^{k,m} \geq x_{1,UB}^{k,m}$, is rejected. This condition occurs when the overall composition for the lower bound, $x_{1,UB}^m$ is higher than the overall composition for the upper bound, $x_{1,UB}^m$. As an example, consider that the composition range for property 1 is within the range 0.5–0.7 while the composition range for property 2 is within the

range 0.3–0.4. According to Equations (5) and (6), $x_{1,LB}^m$ is 0.5 while $x_{1,UB}^m$ is 0.4, which shows that this mixture fulfils the above mentioned condition. Thus, this type of mixture is rejected.

(ii) Level 3: Non-linear constraints. At the end of Level 2, binary mixture candidates (solvent 1, S1 mix with solvent 2, S2) with their composition boundaries have been determined. At this level, non-linear constraints are applied for further screening of the solvent mixtures.

Step 3.1: Solve non-linear models with x_i as input. For this step, non-linear mixture properties, $\zeta^{k,m}$, for the remaining binary mixtures are considered. These non-linear models are solved by using the overall composition range ($x_{1,LB}^m < x_1^m < x_{1,UB}^m$) as input. As a result, new composition ranges that satisfy the non-linear constraints are obtained.

Rule 3: All mixtures for which the calculated property values do not match the non-linear property constraints are rejected. Then the remaining solvent mixtures will be evaluated in Level 4.

(iii) Level 4: Phase stability constraints. At the fourth level, the stability analysis is performed, where the input data used are UNIFAC-LLE group representations (Magnussen *et al.*, 1981) of solvent mixture and operating temperature used in the herbal extraction process.

Step 4.1: Solve the stability routine for each mixture. The stability for each solvent mixture obtained after Level 3 is analysed using a developed stability test. The stability analysis provides the information whether the binary mixtures are stable or not. Stable means that the binary mixtures will not separate (miscible), while unstable means the binary mixtures will separate (either partially miscible or immiscible) at the tested temperature. The stability test is based on the trend of the Gibbs energy and its second derivatives as a function of composition. The Gibbs energy of mixing, ΔG^{mix} is calculated as follows:

$$\frac{\Delta G^{mix}}{RT} = \frac{G^E}{RT} + \sum_{i=1}^{NC} x_i \cdot \ln x_i \quad (7)$$

where G^E is the excess Gibbs energy of mixing, which is calculated from:

$$\frac{G^E}{RT} = \sum_{i=1}^{NC} x_i \cdot \ln \gamma_i \quad (8)$$

Rule 4: Mixtures showing phase separation at x_i are rejected. According to Smith *et al.* (2005), stable mixtures must satisfy these two rules:

- $\frac{\Delta G^{mix}}{RT} < 0$
- $\partial^2 \left(\frac{\Delta G^{mix}}{RT} \right) / \partial x_1^2 > 0$

All solvent mixtures that satisfy both rules will be evaluated at the next level.

(iv) Level 5: Cost calculation. The goal of this algorithm is to obtain the lowest cost of solvent mixtures that satisfy all listed constraints. It is in order to make sure that the solvent properties are within the property constraints as well as no solvent mixture separations occur at the set temperature. At this level, the inputs are solvent mixtures composition as well as pure solvent costs, which are obtained from (ICIS, 2014).

Step 5.1: Identification of the mixture that minimizes the cost. Cost calculations for all stable binary solvent mixtures are performed using linear mixing rules (Equation 1). The composition of solvents resulting from Step 2.2 will be used and substituted in this equation.

2.1.4 Task 4: Performance evaluation

The objective of this task is to evaluate the performance of the solvent blend candidates obtained from the previous step. The performance can be evaluated in terms of extraction capability or profit, or both. Even though a solvent blend may satisfy all of the target properties, this does not guarantee that it is capable of extracting a high yield of phytochemicals. This is because many factors affect the extraction process. It not only depends on the solvents but also the process conditions, size of raw materials (herbs) and mixing process. Thus, a performance evaluation algorithm is developed to solve this problem. The goal of Task 4 is to find the binary solvent mixtures that are able to extract the maximum amount of target herbal phytochemicals as well as return a high profit when the phytochemicals are sold.

Task 4.1 Run performance evaluation algorithm. Figure 6 shows the performance evaluation algorithm used in this study.

The input for this algorithm is the result from the solvent design algorithm (solvent blends with specified composition) together with the extraction temperature and solvent to herb ratio (from the knowledge base). Then, the composition of targeted phytochemicals is computed using a solid–liquid equilibrium, SLE, model (Performance 1). The output of Performance 1 is a list of solvent blends and the extraction yield. Then, the user can choose to continue to the next performance (Performance 2) or terminate the performance evaluation with Performance 1 only. Performance 2 calculates the profit if the phytochemicals are sold by considering the raw material and solvent costs only (the utility cost is not included). The output of this performance is a list of solvent, extraction yield and profit.

Step P1: Identification of mixtures that can extract a target phytochemical. The phytochemicals composition, x_i , is calculated using the SLE model (Smith *et al.*, 2005) as follows:

$$x_i \gamma_i = \exp \frac{\Delta H_i}{RT_m} \left(\frac{T - T_{mi}}{T} \right) \quad (9)$$

where x_i is the solubility of the phytochemicals expressed as mole fraction, γ_i is the phytochemical's activity coefficient in the solution, T_{mi} is the phytochemical's melting temperature, ΔH_i is the phytochemical's heat of fusion and T is the extraction operating temperature.

By using this model, the amount of phytochemicals extracted using the listed solvent blends can be estimated.

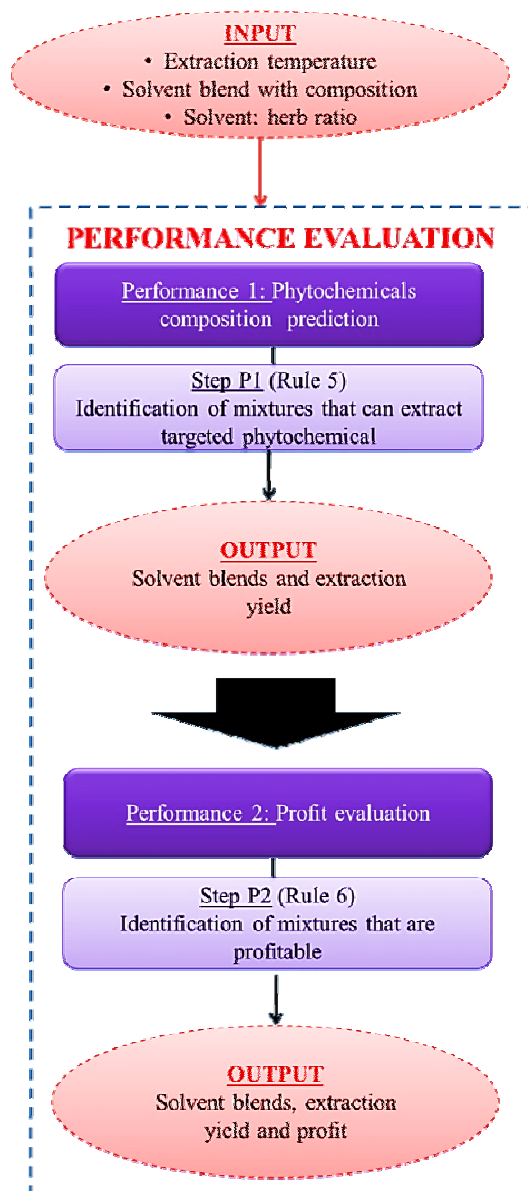


Figure 6: Performance evaluation algorithm

El-Sherbiny *et al.* (2003) mentioned that for dried herb extracts with 80% ethanol (vol/vol), the herb-to-extract ratio, r , is 12:1 for a 100% native extract. Native extract means the material consisting only of components present in the original plant or formed during the extraction process, excluding any excipients or other added substances (OCM, 2011). The herb-to-extract ratio, r , is taken into account while the phytochemical's yield (in g) is computed as shown in Equation (10),

$$Y_p = x_i \times r \times M_{RM} \quad (10)$$

where Y_p is the predicted extraction yield in g, x_i is the predicted extraction yield in wt/wt, and M_{RM} is weight of herbal raw material in g.

Rule 5: Solvent mixtures that give low target phytochemicals composition range (less than 5 wt/wt) are rejected.

Step P2: Identification of mixtures that are profitable

The aim of Performance 2 is to evaluate the profit for solvent blends that can extract a high amount of phytochemical. The profit calculation is shown in Equation (11),

$$P = C_{\text{phyt}} - C_{\text{RW}} \quad (11)$$

$$C_{\text{RW}} = \left(\frac{C_{\text{solv}}}{v \times \rho_{\text{solv}}} + C_{\text{RHerb}} \right) \quad (12)$$

where P is profit (RM/g), C_{phyt} is the phytochemical price (RM/g), C_{RW} is the raw material cost (RM/g), C_{solv} is the solvent cost (RM/g), v is the volume of solvent (cm^3), ρ_{solv} is solvent density (g/cm^3) and C_{RHerb} is the raw herb cost (RM/g).

Rule 6: Solvent blends that give a negative profit if the phytochemicals are sold are rejected.

Task 4.2 Rank solvent blend candidates. The last task in the model-based stage is to rank the solvent blends (result from solvent design and performance evaluation algorithms) according to the selection criteria: solvent cost, profit and predicted phytochemicals composition.

2.1.5 Task 5: Experimental validation

This task is under the experimental verification stage where it is divided into two tasks, namely, herbal extraction and property check. For herbal extraction, the extraction process is performed using the selected solvent blends resulting from the performance evaluation (Task 4).

Task 5.1: Property check. Task 5.1 is needed to verify the model used with the experimental result for the selected property in this study. In Task 2, property model identification is used to search the model based on the targeted property set in Task 1. The calculation of the property is only based on the model, which might have some error and need experimental validation. The percentage error formula is used to determine the precision of the prediction model. Equation (13) shows the calculation of percentage error.

$$\% \text{ Error} = \frac{\text{Experimental Value} - \text{Predicted Value}}{\text{Predicted Value}} \quad (13)$$

A percentage error value of less than 10% will be accepted, while the rest will be rejected. The amendment must be applied to the rejected value where the new model identification for the unsatisfied property is conducted (Task 2).

Task 5.2: Herbal extraction. In Task 4, performance evaluation is performed, where this task only uses the SLE model to calculate the composition of herbal phytochemicals. Thus, this task is very important to compare the model used with real experimental work.

2.2 Work-flow diagram for the solvent design in herbal extraction

The methodology for solvent design in extracting phytochemicals from herbs in combination with the solvent design algorithm can be simplified. In the solvent design

algorithm, every input solvent in each level has two possible results. They are either rejected or continued onto the next level. These results are influenced by the constraints that have been set in the problem definition, Task 1.3. However, only all accepted binary mixture solvents at Level 4 would be considered in the solvent cost calculation at Level 5. Next is the optimal search, which involves the ranking of solvent candidates according to ascending cost. Then, performance evaluation is performed to find the solvent blend that can extract a high level of phytochemicals as well as give a profit when the phytochemicals are sold. Lastly, the experimental validation is performed to prove that the selected solvent blends can be used to extract phytochemicals. The selected properties in this study are also tested to ensure that suitable property models are used. Figure 7 shows the simplified work-flow diagram for the solvent design in herbal extraction.

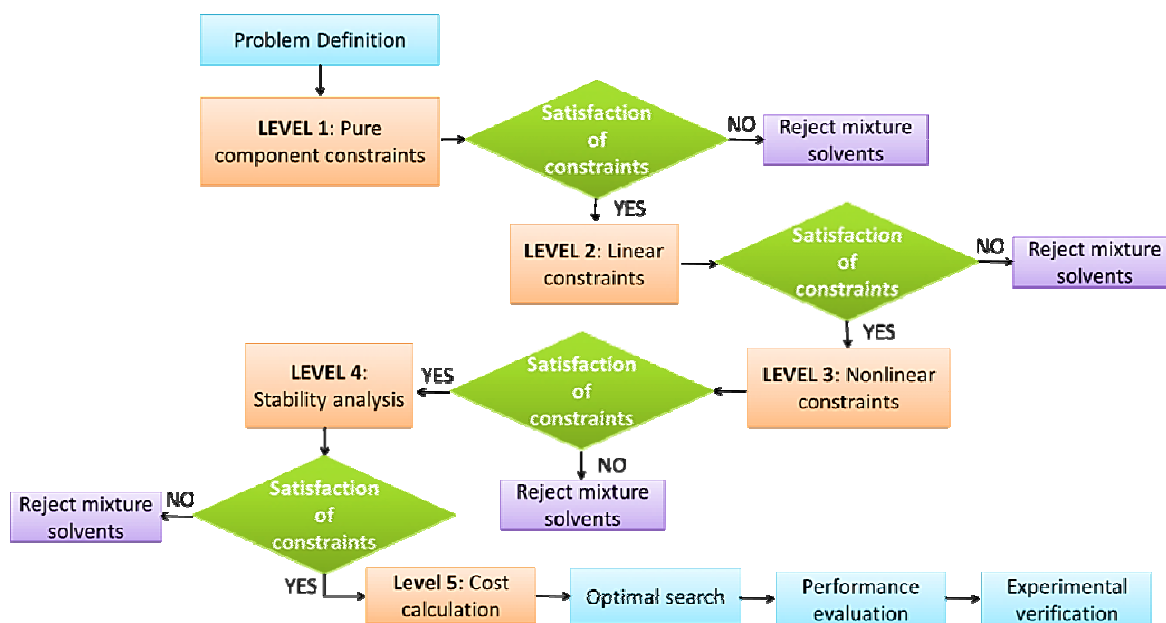


Figure 7: Work-flow diagram for the solvent design in herbal extraction

Table 3 provides a list of the methods and tools used in the solvent design for herbal phytochemicals methodology.

Task 1 uses a knowledge base (any information about herbal extraction processes, the effect of solvent properties on the extraction process, the phytochemical behaviour that could be related to the solvent behaviour, what the experimentally based study considered in the extraction process, how to relate these behaviours to the properties), Task 2 uses a property model library, Task 3 employs a solvent database and solvent design algorithm, Task 4 utilizes a performance evaluation algorithm together with property models library. Task 5 applies experimental procedures and equipment to validate the model-based calculation.

Table 3: Methods and tools used in the solvent design for extraction of herbal phytochemicals

Task	Description	Methods & Tools
Task 1	Problem definition	
	1.1 Understand user needs	- Knowledge base ¹
	1.2 Translate the needs into target properties	- Knowledge base ¹
	1.3 Define constraints on the target properties	- Knowledge base/ toxicity/ phytochemicals consideration
Task 2	Property Model Identification	
	2.1 Retrieve models from the library	- Property models library ²
Task 3	Design Solvent Blend	
	3.1 Choose input data	- Pure solvents data
	3.2 Run solvent design algorithm	- Solvent design algorithm/ stability analysis/solvent cost
	3.3 Rank blends candidates	- Optimal search
Task 4	Performance Evaluation	
	4.1 Phytochemicals composition prediction	- Property models library ² - Performance evaluation algorithm
Task 5	Experimental Validation	
	5.1 Herbal extraction	- Experimental procedure and equipment
	5.2 Property check	- Experimental procedure and equipment

¹ Information from patents and literature.

² Library containing property models for pure components and mixtures, and model for phytochemical composition calculation.

3.0BASE CASE STUDY

The methodology and its implementation are highlighted through a base case study: design of a solvent blend for extracting kaempferol from Kacip Fatimah herb. However, the results that are obtained are only for the model-based design stage.

3.1 Case study: Solvent blends for extracting kaempferol from kacip fatimah herb

The aim of this case study is to design a solvent blend that can maximise the extraction yield of kaempferol, which is one of the main phytochemicals in Kacip Fatimah herb. The solvent blend formulation is considered for non-consumable phytochemicals product used for the conventional extraction and the temperature considered is 90 °C. The result is then compared with the experimental data obtained from Karimi *et al.* (2011). Thirty solvents consisting of alcohols, hydrocarbons, ethers and esters were used as solvent input data, as shown in Table 4.

Table 4: Solvents list with their respective properties

No	Solvents	T _m , K	T _b , K	ΔH _{fus} , kJ/mol	Log K _{ow}	δ, MPa ^{1/2}	-log LC ₅₀ (mol/L)	μ, cP	ρ, g/cm ³
S1	Methanol	196.18	273.26	3.64	-0.23	22.03	2.51	1.32	1.49
S2	Water	253.15	373.15	6.01	0.00	47.84	1.00	0.89	1.00
S3	Acetone	204.41	308.29	6.92	0.32	18.32	-11.12	0.30	0.69
S4	Chloroform	217.55	347.04	9.16	1.73	19.16	-2.21	0.71	1.27
S5	Ethanol	205.36	315.30	6.28	0.15	27.56	14.17	1.63	0.88
S6	Ethyl acetate	187.56	344.07	9.40	0.76	18.18	-2.38	0.43	0.88

Table 4: (continued)

No	Solvents	T _m , K	T _b , K	ΔH _{fus} , kJ/mol	Log K _{ow}	δ, MPa ^{1/2}	-log LC ₅₀ (mol/L)	μ, cP	ρ, g/cm ³
S7	Diethyl ether	112.73	294.38	8.04	1.38	15.28	-3.62	0.29	0.70
S8	Hexane	143.25	338.45	11.07	2.86	14.88	-2.08	0.30	0.74
S9	1-propanol	263.23	362.05	13.30	-1.19	35.57	-2.71	3.42	0.35
S10	Formic acid	214.01	350.66	8.92	0.60	26.37	-8.78	2.02	1.58
S11	Acetic acid	305.74	395.30	9.55	-0.08	17.41	-2.18	1.12	0.93
S12	Propionic acid	310.18	420.56	12.18	0.37	19.19	-2.15	1.39	1.15
S13	Ethylene glycol	275.93	422.54	11.74	-0.78	33.79	0.66	20.83	0.77
S14	1,3-propanediol	288.36	440.00	14.41	-0.30	62.68	2.88	25.76	1.28
S15	Glycerin	291.35	504.45	16.35	-1.44	68.11	0.02	287.18	1.27
S16	Benzene	197.70	361.79	8.88	1.84	18.54	-1.96	0.58	1.06
S17	Cyclohexane	191.31	357.81	3.61	1.59	16.78	-2.34	0.71	0.94
S18	Toluene	208.12	383.75	9.90	2.27	18.04	-1.85	0.60	0.83
S19	Nitrobenzene	298.19	483.72	14.51	1.77	22.35	-1.89	1.96	1.17
S20	Pentane	129.03	300.93	8.43	2.41	14.35	-2.44	0.24	0.69
S21	Cyclopropane	89.11	203.56	0.40	1.06	14.23	-7.87	0.84	0.37
S22	Butyraldehyde	229.77	326.56	12.00	0.79	17.67	-1.68	0.39	1.13
S23	Acetaldehyde	198.70	273.00	10.18	0.17	18.19	-2.00	0.33	0.48
S24	n-Propionaldehyde	207.73	315.09	12.82	0.62	19.48	-1.92	0.41	1.02
S25	n-Butyraldehyde	216.23	350.47	15.46	1.07	17.97	-1.86	0.51	0.75
S26	Isobutyraldehyde	229.77	326.56	12.00	0.79	17.67	-1.68	0.39	0.77
S27	1-Heptanal	239.16	431.77	23.37	2.42	19.66	-1.70	0.96	0.87
S28	1-Hexanal	231.91	407.83	20.73	1.97	19.72	-1.75	0.78	0.85
S29	1-Octanal	246.07	453.39	26.01	2.87	19.69	-1.66	1.19	0.90
S30	Isopropyl alcohol	208.55	329.32	5.23	0.48	25.45	-14.67	1.45	0.91

The properties values are taken from the experimental results. However, if there is no experimental result available, the models in Table 5 were used to predict the desired properties. Meanwhile, the targeted phytochemical in Kacip Fatimah (kaempferol) is used as the phytochemical input data. In this paper, the case study is solved using the systematic methodology illustrated in Figure 1.

Table 5: List of pure property models used in this work

Property of pure solvent	Model	Reference
Partition coefficient, Log K _{ow}	Marrero and Gani GC method	Marrero and Gani (2002)
Boiling point, T _b	Marrero and Gani GC method	Marrero and Gani (2001)
Solubility parameter, δ	Marrero and Gani GC method	Marrero and Gani (2001)
		Mohammad Azmin <i>et al.</i> (2014)
Toxicity, LC ₅₀	GC method	Gao <i>et al.</i> (1992)
Viscosity, μ	GC method	Cao <i>et al.</i> (1993)
Density, ρ	Modified Rackett equation	Spencer and Danner (1972)

*GC method is group contribution method

3.1.1 Task 1: Problem definition

Task 1.1 Identify needs. In this task, an understanding of the consumer needs would be the performance criterion. For herbal extraction, the solvents must have all of the following performance criteria, namely: can effectively extract the selected phytochemicals from the herb, can be removed from the crude extract mixture (so that

the pure crude extract can be obtained), have low toxicity, must be miscible with each other and with the phytochemical, stable, low price and good solvent appearance.

Task 1.2 Translate needs into target properties. According to the knowledge base, the solvent needs are translated to the target properties. Therefore, the target properties affecting the performance criteria are listed in Table 6.

Table 6: Translation of the performance criteria into target properties

Solvents Mixture Performance Criteria	Target Properties
Effectively extract the selected phytochemicals	$\log p, \delta$
Can be removed from phytochemicals crude extract (after extraction)	μ, T_b
Low toxicity	LC_{50}
Miscible with each other and stable	$\delta, \Delta G^{\text{mix}}$
Low price	ρ, C
Good solvent appearance	μ

Task 1.3 Define constraints on the target properties. The target value for all of the listed properties refers to the common existing solvent used in the herbal extraction process, properties of the target phytochemical and literature study. The target values for each property are set as listed in Table 7.

Table 7: Target property constraints

Property	Target property value	
	Solvent constraints	Phytochemical constraints
Partition coefficient	Log K_{ow} (follow phytochemicals) target	$-0.3 \leq \text{Log } K_{ow} \leq 4.44$
Boiling point	$T_b \leq 728.4 \text{ K}$	-
Toxicity parameter	$-2.5 \leq -\log LC_{50} \leq 2.5$	-
Stability	ΔG^{mix}	-
Viscosity	$0.20 \text{ cP} \leq \mu \leq 1.24 \text{ cP}$	-
Density	$1.0 \text{ g/cm}^3 \leq \rho \leq 1.5 \text{ g/cm}^3$	-
Price	C	-
Solubility parameter	δ (follow phytochemicals) target	$16 \leq \delta \leq 48 \text{ Mpa}^{1/2}$

3.1.2 Task 2: Property model identification

Task 2.1: Retrieve models from the library. The target properties, partition coefficient ($\log K_{ow}$), toxicity parameter (LC_{50}), solubility parameter (δ), viscosity (μ), density (ρ) and cost (C) are estimated using linear mixing rules (Equation 1) while the others are predicted by using non-linear models, as listed in the property models library (Table 8).

Table 8: List of mixture property models used in this work

Target property	Model	Reference
Partition coefficient, Log K_{ow}	Linear mixing rule	
Toxicity, LC_{50}	Linear mixing rule	
Solubility parameter, δ	Linear mixing rule	
Cost	Linear mixing rule	

Boiling point, T_b	$R = 1 - \sum_{i=1}^N \frac{x_i \gamma_i(x_i, T_b) P_i^\circ(T_i)}{P} = 0$	Klein <i>et al.</i> (1992)
Table 8: (continued)		
Target property	Model	Reference
Stability, ΔG^{mix}	$\frac{\Delta G^{\text{mix}}}{RT} = \frac{G^E}{RT} + \sum_{i=1}^{NC} n_i \cdot \ln(x_i),$ $\frac{G^E}{RT} = \sum_{i=1}^{NC} \chi_i \cdot \ln(\gamma_i)$	Pinal <i>et al.</i> (1991)
Viscosity, μ	$\mu_B = \exp \sum_i x_i \ln \mu_i$	Mehrotra <i>et al.</i> (1996)
Density, ρ	$\rho_B = \sum_i v_i \rho_i(T)$	Yunus <i>et al.</i> (2011)

3.1.3 Task 3: Design solvent blend

Task 3.1: Choose input data. Two sets of input data were used in this case study, the solvents and phytochemicals databases. The solvents input data consist of alcohols, hydrocarbons, ethers and esters, while four main phytochemicals in Kacip Fatimah: kaempferol, myricetin, quercetin and rutin were used as phytochemical input data. Input data means the data containing all of the considered pure properties of both solvents and phytochemicals.

Task 3.2: Run solvent design algorithm. The solvent design algorithm is run following all listed levels.

(i) Level 1: Pure component constraints

Step 1.1: Pure component properties of the solvents in the input data and target phytochemicals are compared with respect to the target values. The aim for this level is to obtain a list of pure solvents that match the phytochemical target property values. For this step, two properties, the solubility parameter (δ) and partition coefficient ($\log p$) are considered. These properties have an interrelation between solvent and phytochemicals that affects the extraction process efficiency while the other properties are used for the solvent selection only to ensure their safety and compatibility with the extraction process.

After considering all of the constraints set in this level, 119 binary solvent combinations out of 870 possible total combinations of binary solvents satisfied all of the constraints. These binary solvents combination will be further screened in level 2.

(ii) Level 2: Linear constraints

Step 2.1: In this level, only properties that satisfy the linear mixing rule (Equation 1) as shown in the property library (Table 8) for binary solvent properties calculation will be considered. In this case, toxicity, density and viscosity are taken into account. The result for this step is the composition range for each binary mixture.

Step 2.2: In this step, the overall composition range for each mixture is identified. This step gives the feasible mixtures with their upper and lower bounds of composition range. After performing Level 2, 36 binary solvents were left and will be considered in Level 3.

(iii) Level 3: Non-linear constraints

Step 3.1: Level 3 is only for the properties that satisfy the non-linear model for binary properties calculation as shown in the property library (Table 8). In this case, only one property, boiling point, was applied to the non-linear model. The boiling point model was solved with x_i (composition boundary) as input.

Step 3.2: Mixtures that do not match non-linear constraints are rejected. After performing Level 3, only 12 binary solvents still remain.

(iv) Level 4: Stability analysis

Step 4.1: The remaining binary solvents are then described based on the UNIFAC-LLE group representation and the activity coefficients are calculated at the temperature of 90 °C (operating temperature for the extraction of kaempferol from Kacip Fatimah) and the stability tests are performed for all of the remaining candidates.

Step 4.2: All mixtures that show a phase separation at the composition boundary, x_i , are rejected. However, in this case study, none are rejected because none of the 12 binary solvents showed a phase separation. All 12 solvent mixtures with their properties considered in this study are shown in Table 9.

Table 9: Binary solvents that satisfy Level 4 with their properties

Solvent mixtures	x_i	δ , Mpa ^{1/2}	Log K_{ow}	T_m	$-\log LC_{50}$	ρ , g/cm ³	μ , cP	T_b , K
M:IB	0.92	21.67	-0.15	198.87	2.17	1.46	1.20	338.15
M:PP	0.91	21.79	-0.15	197.22	2.11	1.45	1.19	321.15
M:W	0.8	27.18	-0.18	227.37	2.21	1.39	1.22	340.65
M:EA	0.35	19.52	0.41	190.58	-0.67	1.09	0.64	333.65
M:AA	0.4	19.25	-0.04	261.92	-0.30	1.15	1.20	346.15
PD:PP	0.03	20.78	0.59	210.15	-1.78	1.03	0.46	321.15
PD:IB	0.03	19.02	0.76	231.53	-1.54	1.14	0.44	338.15
M:PA	0.02	19.25	0.36	307.90	-2.06	1.16	1.39	336.65
G:B	0.01	19.04	1.81	198.64	-1.94	1.06	0.62	334.15
PD:P	0.03	16.09	2.38	136.39	-2.33	0.72	0.27	321.15
PD:B	0.03	19.86	1.78	200.42	-1.81	1.06	0.65	322.15
W:PA	0.7	39.25	0.11	270.26	0.06	1.04	1.02	340.14

** M: methanol, IB:isobutyraldehyde, PP: *n*-propionaldehyde, W: water, EA: ethyl acetate, AA: acetic acid, PD: 1,3-propanediol, PA: propionic acid, B: benzene, G: glycerol

The reduction in the numbers of solvent mixtures in all four levels is illustrated in Figure 8.

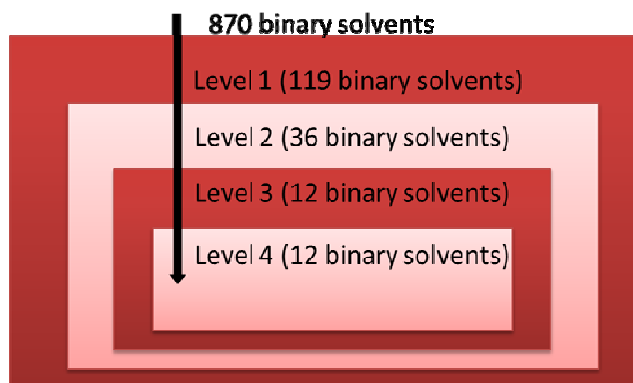


Figure 8: Number of solvent blends reduced after screening using the developed methodology

(v) Level 5: Cost calculation

Step 5.1: Identification of the mixture that minimizes the cost is performed for all 12 binary solvents that satisfy Levels 1 to 4. Table 10 shows the cost in RM per ml of single solvent. Table 11 shows the solvent mixture price for all 12 binary solvents at their respective composition.

Table 10: Price considered in this study

Solvent/phytochemicals	Price (RM/ml)
Methanol, M	1.78×10^{-3}
Water, W	2.70×10^{-6}
Isobutyraldehyde, IB	1.39×10^{-1}
n-propionedehyde, PP	3.16×10^{-1}
Ethyl acetate, EA	3.54×10^{-3}
Acetic acid, AA	5.24×10^{-3}
Priopionic acid, PA	3.15×10^{-1}
1,3-propylene glycol, PD	4.18×10^{-3}
Benzene, G	1.25×10^{-3}
Glycerol, B	1.26×10^{-3}

Table 11: Solvent mixture price

Solvent Mixture	x_1	Price, RM/ml
M:IB	0.92	0.0123
M:PP	0.91	0.0291
M:W	0.8	0.0014
M:EA	0.35	0.0028
M:AA	0.4	0.0037
PG:PP	0.03	0.2964
PG:IB	0.03	0.1303
M:PA	0.02	0.2994
G:B	0.01	0.3025
W:PG	0.01	0.0040
PG:B	0.03	0.0034
W:PA	0.7	0.0916

3.1.4 Task 4: performance evaluation

The performance evaluation is divided into two main tasks. The first task is the prediction of phytochemical composition while the second task is profit evaluation. Both tasks are important as they will evaluate which binary solvents would extract the highest phytochemical composition with the highest profit.

Task 4.1: Run Performance Evaluation Algorithm

(i) Performance 1: Phytochemical composition prediction. This step is performed to identify which solvent mixtures (from the 12 binary solvents remaining after Task 3) can extract the target phytochemical. Original UNIFAC and Modified UNIFAC (Dortmund) models were used to estimate the activity coefficients as formulated in Equation (9). The estimated kaempferol composition was verified with experimental data from Karimi *et al.* (2011). The original UNIFAC gave 85.12% error while the Modified UNIFAC (Dortmund) gave 21.46% error when compared with the experimental data. Thus, this work used the Modified UNIFAC (Dortmund) to calculate the kaempferol composition in Kacip Fatimah herb based on the lower error obtained. By using Equation (9), the compositions of phytochemicals that can be extracted using the 12 binary solvent combinations are predicted and the result is shown in Figure 9.



Figure 9: Predicted yield of kaempferol composition extracted by solvent blends in rank

The highest yield of kaempferol composition is 0.4202 wt/wt for M:IB, while the lowest is 0.0 wt/wt for W:PA. 0.0 wt/wt means that this solvent mixture could not extract kaempferol even though all of the solvent properties are within the target property constraints. In this design, a composition value with less than 5 wt/wt is considered very low and can be neglected. As an example, for G:B solvent mixture, it can only extract 0.0011 wt/wt of kaempferol and the M_{RM} used is 2 g. After substituting this value into Equation (10), the yield obtained is only 0.000183 g of kaempferol. If the value for

0.4202 wt/wt (M:IB) and M_{RM} use of 2 g are substituted into Equation (10), the result obtained is 0.07 g. The value calculated for G:B is very small compared with the M:IB mixture. Therefore, seven solvent mixtures that gave less than 5 wt/wt of yield are rejected, namely, PG:PP, PG:IB, M:PA, G:B, W:PG, PG:B and W:PA. Thus, only five binary solvents remain for the next performance (Performance 2). The ranking for the top five binary solvents according to the kaempferol composition yield is M:IB followed by M:PP, M:W, M:EA and M:AA.

The cost of the top five binary solvents (calculated in Step 5.1, Solvent Design Algorithm) and the yield of kaempferol that could be extracted is plotted in Figure 10. From this figure, it is seen that M:W gives the highest predicted kaempferol yield at the lowest cost. Nevertheless, the next task (Performance 2) will reveal the profit that might be obtained if this phytochemical was sold.

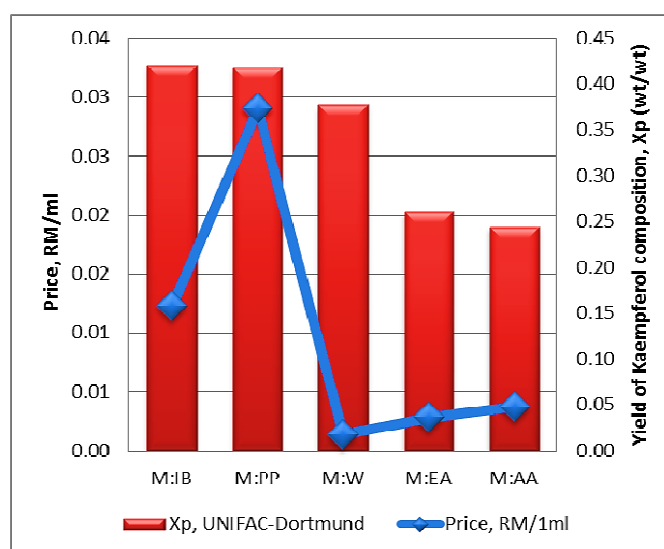


Figure 10: Cost evaluation for extraction of Kacip Fatimah herb

(ii) Performance 2: Profit evaluation. The profit evaluation objective is to compute the obtained profit by considering all raw materials cost (solvents and Kacip Fatimah herb) and selling price for the extracted crude oil. The selling price for kaempferol with 90% purity is RM 8.58/mg. Input data used in this calculation are raw Kacip Fatimah and solvent prices, which were obtained from Herbs (2014) and (ICIS (2014)), respectively.

As shown in Figure 11, extraction of 1 mg of kaempferol using M:PP gives the highest profit (RM 10.24) compared with the others with moderate raw material cost (RM 0.0027). Then, it is followed by M:IB with a profit of RM 10.00 and the raw material cost is RM 0.0026. By using M:W, it can give a profit of RM 8.58 per mg of kaempferol, which is the same profit as for M:EA and M:AA. However, the raw material cost for M:W is RM 0.0025, M:EA is RM 0.0042 and M:AA is RM 0.0045. This profit and loss calculation excludes manpower, utility and electricity costs. From Figure 11, it can be concluded that M:PP not only gives the highest profit among the others but it also can extract the second highest amount of kaempferol, as also shown in Figure 10.

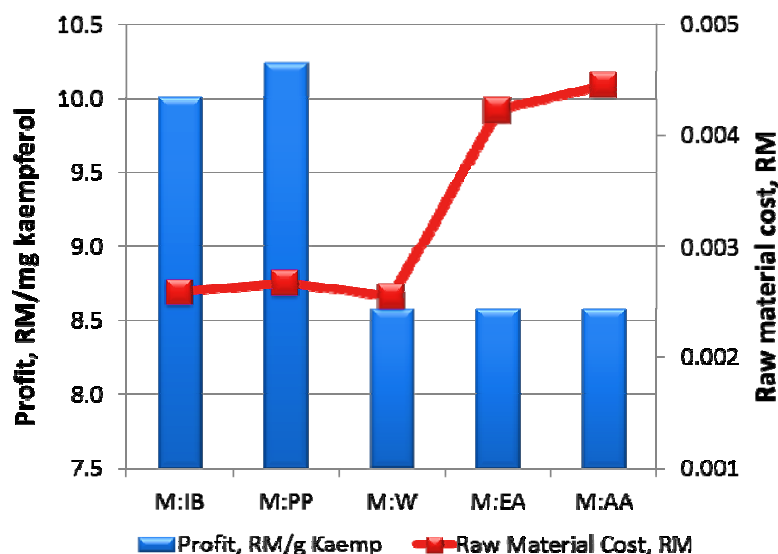


Figure 11: Cost evaluation for extraction of Kacip Fatimah herb.

Task 4.2: Rank solvent blends candidates

The purpose of this task is to rank the binary solvents that can give the highest profit (which have the lowest solvent costs and have the potential to extract the selected phytochemicals). All five binary solvents are ranked according to descending profit and phytochemicals composition, as illustrated in Figure 12.

Figure 12 shows that the M:IB mixture has the highest extraction yield. However, the profit for this mixture is ranked second. On the other hand, M:PP mixture has the highest profit but is ranked second in terms of extraction yield. High profit in this study means, the mixture can extract high yield of phytochemicals with low raw material cost. Wu (2012) stated that the low-cost manufactured product would always be chosen by an industrial company when the product is to be sold. Thus, the profit is chosen as the first priority as compared with the extraction yield. In this study, the M:PP mixture is chosen as the mixture that can produce a high extraction yield as well as profit.

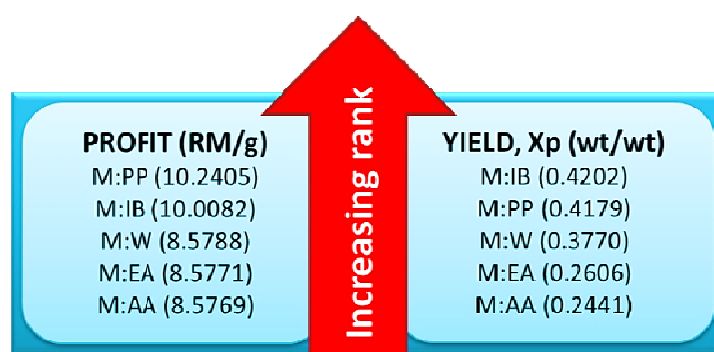


Figure 12: The rank for the profit and extraction yield

4.0 CONCLUSIONS

A systematic methodology for the design of blended solvent for extracting phytochemicals from herbs has been developed and was tested on the extraction of kaempferol from Kacip Fatimah herb. A decomposition method has been applied to solve the solvent blending problem for extraction of phytochemicals, where the

objectives are quickly to screen out a large number of alternatives and to reduce the search space at each hierarchical step. The methodology can be used to design blended solvent for extracting phytochemicals from any herb where the scope and size of the case study depend on the solvent database available and availability of models. This methodology is described and highlighted for binary solvent mixtures, but can easily be extended to multicomponent mixtures. For future work, this systematic methodology needs to be verified for extraction of different phytochemicals from various herbs as case studies.

ACKNOWLEDGEMENT

This work was supported by the Fundamental Research Grant Scheme, FRGS (Vote number: R.J130000.7809.4F488), Universiti Teknologi Malaysia (UTM), Universiti Malaysia Kelantan (UMK) and the Ministry of Education, Malaysia. This support is gratefully acknowledged.

REFERENCES

- Abdullah, F.; Ling, S. K.; Man, S.; Tan, A. L.; Tan, H. P. and Abdullah, Z. (2012). Characterization and identification of *Labisia pumila* by multi-steps infrared spectroscopy. *Vibrational Spectroscopy*, 62, 200-206.
- Barton, A. F. (1990). *Polymer-liquid interaction parameters and solubility parameters*. Boca Raton Florida, USA: CRC press.
- Cheng, H. C. and Wang, F. S. (2007). Trade-off optimal design of a biocompatible solvent for an extractive fermentation process. *Chemical Engineering Science*, 62, 4316-4324.
- Cheng, H. C. and Wang, F. S. (2008). Optimal process/solvent design for ethanol extractive fermentation with cell recycling. *Biochemical Engineering Journal*, 41, 258-265.
- Cheng, H. C. and Wang, F. S. (2010). Computer-aided biocompatible solvent design for an integrated extractive fermentation–separation process. *Chemical Engineering Journal*, 162, 809-820.
- Choi, H.-K.; Kim, D.-H.; Kim, J. W.; Ngadiran, S.; Sarmidi, M. R. and Park, C. S. (2010). *Labisia pumila* extract protects skin cells from photoaging caused by UVB irradiation. *Journal of Bioscience and Bioengineering*, 109, 291-296.
- Conte, E.; Gani, R.; Cheng, Y. S. and Ng, K. M. (2012). Design of formulated products: Experimental component. *AIChE Journal*, 58, 173-189.
- Conte, E.; Gani, R. and Ng, K. M. (2011). Design of formulated products: a systematic methodology. *AIChE Journal*, 57, 2431-2449.
- Damartzis, T.; Kouneli, A.; Papadopoulos, A. I.; Seferlis, P.; Dimitriadis, G. and Vlachopoulos, G. (2014). Optimal design of solvent based post combustion CO₂ capture processes in quicklime plants. *Chemical Engineering*, 39.
- Diallo, D.; Hveem, B.; Mahmoud, M. A.; Berge, G.; Paulsen, B. S. and Maiga, A. (1999). An ethnobotanical survey of herbal drugs of Gourma district, Mali. *Pharmaceutical Biology*, 37, 80-91.
- El-Sherbiny, D. A.; Khalifa, A. E.; Attia, A. S. and Eldenshary, E. E. S. (2003). *Hypericum perforatum* extract demonstrates antioxidant properties against elevated rat brain oxidative status induced by amnestic dose of scopolamine. *Pharmacology Biochemistry and Behavior*, 76, 525-533.

- Folić, M.; Adjiman, C. S. and Pistikopoulos, E. N. (2005). A computer-aided methodology for optimal solvent design for reactions with experimental verification. *Computer Aided Chemical Engineering*, 20, 1651-1656.
- Gani, R. (2004). Computer-aided methods and tools for chemical product design. *Chemical Engineering Research and Design*, 82, 1494-1504.
- Gani, R. and Brignole, E. (1983). Molecular design of solvents for liquid extraction based on UNIFAC. *Fluid Phase Equilibria*, 13, 331-340.
- Garg, S. and Achenie, L. E. (2001). Mathematical programming assisted drug design for nonclassical antifolates. *Biotechnology progress*, 17, 412-418.
- Herbs, D. J. (2014). *Pembekal kacang fatimah* [Online]. Malaysia. Available: <http://www.delimajelita.com/2009/12/serbuk-kacip-fatimah.html> [Accessed 24 November 2014].
- Hill, M. (2009). Chemical product engineering—the third paradigm. *Computers & Chemical Engineering*, 33, 947-953.
- Icis. (2014). *Indicative Chemical Prices A-Z* [Online]. New York: ICIS. Available: <http://www.icis.com/chemicals/channel-info-chemicals-a-z/> [Accessed 10 October 2014].
- Karimi, E.; Jaafar, H. Z. and Ahmad, S. (2011). Phytochemical analysis and antimicrobial activities of methanolic extracts of leaf, stem and root from different varieties of *Labisia pumila* Benth. *Molecules*, 16, 4438-4450.
- Karimi, E.; Jaafar, H. Z. and Ahmad, S. (2013). Antifungal, anti-inflammatory and cytotoxicity activities of three varieties of *labisia pumila* benth: from microwave obtained extracts. *BMC complementary and alternative medicine*, 13, 20.
- Karunanithi, A. T.; Achenie, L. E. and Gani, R. (2005). A new decomposition-based computer-aided molecular/mixture design methodology for the design of optimal solvents and solvent mixtures. *Industrial & engineering chemistry research*, 44, 4785-4797.
- Karunanithi, A. T.; Achenie, L. E. and Gani, R. (2006). A computer-aided molecular design framework for crystallization solvent design. *Chemical engineering science*, 61, 1247-1260.
- Karunanithi, A. T.; Acquah, C.; Achenie, L. E.; Sithambaram, S. and Suib, S. L. (2009). Solvent design for crystallization of carboxylic acids. *Computers & Chemical Engineering*, 33, 1014-1021.
- Kerton, F. M. and Marriott, R. (2013). *Alternative solvents for green chemistry*. (2nd ed). Cambridge, UK: Royal Society of chemistry.
- Klein, J.; Wu, D. and Gani, R. (1992). Computer aided mixture design with specified property constraints. *Computers & Chemical Engineering*, 16, S229-S236.
- Kumoro, A. C. and Hasan, M. (2008). Extraction of herbal components—the case for supercritical fluid extraction. *Teknik*, 29, 180-183.
- Lee, C.; Choy, K. and Chan, Y. (2014). A knowledge-based ingredient formulation system for chemical product development in the personal care industry. *Computers & Chemical Engineering*, 65, 40-53.
- Liao, M.; Zhao, Y.; Ning, P.; Cao, H.; Wen, H. and Zhang, Y. (2014). Optimal design of solvent blend and its application in coking wastewater treatment process. *Industrial & Engineering Chemistry Research*, 53, 15071-15079.
- Magnussen, T.; Rasmussen, P. and Fredenslund, A. (1981). UNIFAC parameter table for prediction of liquid-liquid equilibriums. *Industrial & Engineering Chemistry Process Design and Development*, 20, 331-339.

- Norhaiza, M.; Maziah, M. and Hakiman, M. (2009). Antioxidative properties of leaf extracts of a popular Malaysian herb, *Labisia pumila*. *J. Med. Plant Res*, 3, 217-223.
- Ocm, M. (2011). Guidance on equivalence of herbal extracts in complementary medicines.
- Papadopoulos, A. I. and Seferlis, P. (2009). *A Framework for Solvent Selection Based on Optimal Separation Process Design and Controllability Properties*. In: Jacek, J. and Jan, T. (eds.) *Computer Aided Chemical Engineering*. (177-181). Elsevier.
- Samudra, A. P. and Sahinidis, N. V. (2013). Optimization-based framework for computer-aided molecular design. *AIChE Journal*, 59, 3686-3701.
- Smith, J. M.; Van Ness, H. C. and Abbott, M. M. (2005). *Introduction to chemical engineering thermodynamics*. Boston: McGraw-Hill; 7th ed.
- Stavrou, M.; Lampe, M.; Bardow, A. and Gross, J. (2014). Continuous Molecular Targeting-Computer-Aided Molecular Design (CoMT-CAMD) for Simultaneous Process and Solvent Design for CO₂ Capture. *Industrial & Engineering Chemistry Research*, 53, 18029-18041.
- Vanderveen, J. R.; Durelle, J. and Jessop, P. G. (2014). Design and evaluation of switchable-hydrophilicity solvents. *Green Chemistry*, 16, 1187-1197.
- Wu, C.-H. (2012). Product-design and pricing strategies with remanufacturing. *European Journal of Operational Research*, 222, 204-215.
- Yunus, N. A.; Gernaey, K. V.; Woodley, J. M. and Gani, R. (2014). A systematic methodology for design of tailor-made blended products. *Computers & Chemical Engineering*, 66, 201-213.
- Cheng, H. C. and Wang, F. S. (2007). Trade-off optimal design of a biocompatible solvent for an extractive fermentation process. *Chemical Engineering Science*, 62, 4316-4324.
- Cheng, H. C. and Wang, F. S. (2008). Optimal process/solvent design for ethanol extractive fermentation with cell recycling. *Biochemical Engineering Journal*, 41, 258-265.
- Cheng, H. C. and Wang, F. S. (2010). Computer-aided biocompatible solvent design for an integrated extractive fermentation-separation process. *Chemical Engineering Journal*, 162, 809-820.
- Conte, E.; Gani, R.; Cheng, Y. S. and Ng, K. M. (2012). Design of formulated products: Experimental component. *AIChE Journal*, 58, 173-189.
- Damartzis, T.; Kouneli, A.; Papadopoulos, A. I.; Seferlis, P.; Dimitriadis, G. and Vlachopoulos, G. (2014). Optimal design of solvent based post combustion CO₂ capture processes in quicklime plants. *Chemical Engineering*, 39.
- Folić, M.; Adjiman, C. S. and Pistikopoulos, E. N. (2005). A computer-aided methodology for optimal solvent design for reactions with experimental verification. *Computer Aided Chemical Engineering*, 20, 1651-1656.
- Karunanithi, A. T.; Achenie, L. E. and Gani, R. (2005). A new decomposition-based computer-aided molecular/mixture design methodology for the design of optimal solvents and solvent mixtures. *Industrial & engineering chemistry research*, 44, 4785-4797.
- Karunanithi, A. T.; Achenie, L. E. and Gani, R. (2006). A computer-aided molecular design framework for crystallization solvent design. *Chemical engineering science*, 61, 1247-1260.

- Liao, M.; Zhao, Y.; Ning, P.; Cao, H.; Wen, H. and Zhang, Y. (2014). Optimal design of solvent blend and its application in coking wastewater treatment process. *Industrial & Engineering Chemistry Research*, 53, 15071-15079.
- Papadopoulos, A. I. and Seferlis, P. (2009). *A Framework for Solvent Selection Based on Optimal Separation Process Design and Controllability Properties*. In: Jacek, J. and Jan, T. (eds.) *Computer Aided Chemical Engineering*. (177-181). Elsevier.
- Stavrou, M.; Lampe, M.; Bardow, A. and Gross, J. (2014). Continuous Molecular Targeting–Computer-Aided Molecular Design (CoMT–CAMD) for Simultaneous Process and Solvent Design for CO₂ Capture. *Industrial & Engineering Chemistry Research*, 53, 18029-18041.
- Vanderveen, J. R.; Durelle, J. and Jessop, P. G. (2014). Design and evaluation of switchable-hydrophilicity solvents. *Green Chemistry*, 16, 1187-1197.