

## Greener Solvent Selection Under Uncertainty

Ki-Joo Kim<sup>1</sup>, Urmila M. Diwekar<sup>2,\*</sup>, and Kevin G. Joback<sup>3</sup>

<sup>1,2</sup> CUSTOM,

Environmental Eng., Carnegie Mellon University, Pittsburgh, PA 15213  
<sup>3</sup>Molecular Knowledge Systems, P.O.Box 10755, Bedford, NH 03110-0755

Solvent selection is an important step in process synthesis, design, or process modification. Computer-aided molecular design (CAMD) approach based on the reverse use of group contribution method provides a promising tool for solvent selection. However, uncertainties inherent in these techniques and associated models are often neglected. This paper presents a new approach to solvent selection under uncertainty. A case study of acetic acid extraction demonstrates the usefulness of this approach to obtain robust decisions.

### Introduction

Solvents are extensively used as process materials (e.g., extracting agent) or process fluids (e.g., CFC) in chemical process industries, pharmaceutical industries, and solvent-based industries such as coating and painting. Since waste solvents are a main source of pollution to air, water, and soil, it is desirable to use reduced amount of solvents and/or environmentally friendly solvents without sacrificing process performance. There are some solvents which must be

eliminated due to environmental and health effects and regulatory requirements. For example, the Montreal protocol bans many chlorinated solvents (1).

Solvent selection, an approach to generate candidate solvents having desirable properties, can help to handle these problems. Several methodologies have been developed for solvent selection over the years (2). First approach uses traditional laboratory synthesis and test methodology to find promising solvents. This method can provide reliable and accurate result, but in many cases this method can not be applied due to cost, safety, and time constraints. Second approach is to search the property database. Although it is the most common and simple method, it is limited by the size and accuracy of the database. Furthermore, these two methods may not provide the best solvent because there are huge number of solvent molecules to be tested or searched. Finally, computer-aided molecular design (CAMD) can automatically generate promising solvents from their building blocks or groups (3,4). This method can generate lists of candidate solvents with reasonable accuracy within moderate time scale. CAMD can also be applied to CFC substituents (4), solvent blend design (5), polymer and drug design (6), and alternative process fluid design (7). However, CAMD is limited by the availability and reliability of the property estimation methods.

All methodologies for solvent selection are exposed to uncertainties that arise from experimental errors, imperfect theories or models and their parameters, improper knowledge or ignorance of systems, and inadequate controls. Although uncertainties can affect the real implementation of selected solvents, few papers in literature have focused on uncertainties. In this study a new CAMD method for solvent selection under uncertainty is presented and is applied to generate greener solvents for acetic acid extraction from water as a case study.

As the ranking or priority for solvent selection in this study is based on the Hansen's solubility parameters, we describe solubility parameters and solvent selection model first. Then several sections are devoted to explain and discuss this new CAMD method, uncertainty quantification, and case study. Finally summary is followed.

### **Hansen's Solubility Parameter and Solvent Selection Model**

The solubility parameter,  $\delta$ , is one of the most important parameters in physical chemistry and thermodynamics of solutions. It can serve as a key parameter for solvent selection, solubility estimation, and the estimation of polymer swelling (8). Though it was originally introduced by Hildebrand and Scott (9), the most common one is Hansen's three-dimensional solubility parameter (10) which is given by:

$$\delta^2 = \delta_d^2 + \delta_p^2 + \delta_h^2 \quad (\text{unit: MPa}) \quad (1)$$

where  $\delta_d$  is the dispersive term,  $\delta_p$  is the polar term, and  $\delta_h$  is the hydrogen bonding term. The solubility parameter ( $\delta$ ) and its three terms ( $\delta_d$ ,  $\delta_p$ , and  $\delta_h$ ) can be determined by semi-empirical methods and are tabulated by Barton (11) for most common liquids.

Miscibility of two liquids A and B depend on the heat of mixing  $\Delta H_{\text{mix}}$ , and  $\Delta H_{\text{mix}}$  in the Hansen theory is given by the following equation.

$$\Delta H_{\text{mix}} = (n_A V_A + n_B V_B) \left[ (\delta_d^A - \delta_d^B)^2 + (\delta_p^A - \delta_p^B)^2 + (\delta_h^A - \delta_h^B)^2 \right] \Phi_1 \Phi_2 \quad (2)$$

where  $n$  is the number of moles,  $V$  is the molar volume, and  $\Phi$  is the volume fraction. When the heat of mixing approaches to zero, two liquids A and B are soluble or miscible with each other. Hence the three solubility parameter terms should be close to minimize the heat of mixing.

Generally speaking, solubility parameters for solvent selection are not as accurate as other property estimation methods such as infinite dilution activity coefficients ( $\gamma^\infty$ ). But solubility parameter method is universal and simple to apply, and thus can be used for guiding and screening candidate solvents with relatively acceptable accuracy.

To replace the current solvent or design new solvents, there are several criteria are used. These are (a) distribution coefficient ( $m$ ), (b) solvent selectivity ( $\beta$ ), (c) solvent loss ( $S_L$ ), (d) physical properties such as boiling point, flash point, density, and viscosity, (e) chemical stability, (f) toxicology, and (g) cost. For extraction process, the final selection of solvents will generally be dominated between  $m$  and  $\beta$ . Distribution coefficient ( $m$ ), a measure of solvent capacity, is defined as:

$$m \propto \left( \frac{r_{UR}}{r_{US}} \right)^2 \frac{MW_R}{MW_S} \quad (3)$$

where R, U, and S are raffinate, solute, and solvent, respectively. MW is molecular weight.  $r_{UR}$  and  $r_{US}$  are Euclidean distance metric between two molecules and are defined as follows:

$$r_{UR} = \left[ (\delta_d^U - \delta_d^R)^2 + (\delta_p^U - \delta_p^R)^2 + (\delta_h^U - \delta_h^R)^2 \right]^{1/2}$$

$$r_{US} = \left[ (\delta_d^U - \delta_d^S)^2 + (\delta_p^U - \delta_p^S)^2 + (\delta_h^U - \delta_h^S)^2 \right]^{1/2} \quad (4)$$

Liquids having similar solubility parameters are soluble or miscible with each other. The distance between solute and solvent ( $r_{US}$ ) should be small while the distance between solute and raffinate ( $r_{UR}$ ) is fixed for a given solute-raffinate system. Thus solvents having smaller  $r_{US}$  can increase  $m$ , and high  $m$  reduces the size of an extracting equipment and the amount of recycling solvent.

Solvent selectivity ( $\beta$ ), the ability of the solvent to selectively dissolve solute, is the ratio between distribution coefficients of solute and raffinate, and defined by:

$$\beta = \frac{m_U}{m_R} \propto \left( \frac{r_{RS}}{r_{US}} \right)^2 \frac{MW_U}{MW_R} \quad (5)$$

where  $r_{RS}$  is defined in a similar way. High  $\beta$  reduces the cost of solute recovery as solvent is highly selective to solute. Solvent loss ( $S_L$ ) can be expressed by the following equation:

$$S_L \propto \left( \frac{1}{r_{SR}} \right)^2 \frac{MW_S}{MW_R} \quad (6)$$

Low  $S_L$  means high selectivity toward solute and determines immiscibility between solvent and raffinate.

Table 1 shows the Hansen's three-dimensional solubility parameters and estimated solvent properties ( $m$ ,  $\beta$ , and  $S_L$ ). In this example, acetic acid is a solute, water is a raffinate, and ethyl acetate is a current solvent. The aim of solvent selection is to generate solvents having better solvent properties than the

**Table 1. The Hansen's solubility parameters and solvent properties.**

Solvent	$\delta_d$	$\delta_p$	$\delta_h$	Properties	Values
Acetic acid	13.9	12.2	18.9	$m$	1.043
Water	12.2	22.8	40.4	$\beta$	35.20
Ethyl acetate	13.4	8.60	8.90	$S_L$	0.0041

current solvent.

### Computer Aided Molecular Design Under Uncertainty

Group contribution method is a forward problem; if we know a molecule, we can estimate its physical, chemical, biological, and health effect properties based on its groups or building blocks (Figure 1a). However, CAMD is a backward problem; if we know desirable properties or regulation limits, we can find molecules that satisfy these properties or limits by constructing groups (Figure 1b). CAMD approach, though not as accurate as experimentation, can generally provide satisfactory results from large scale combinations of groups.

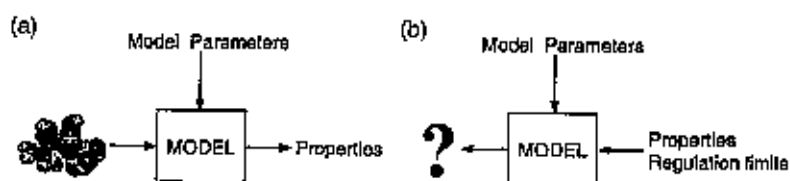


Figure 1. Forward(a) and backward(b) problems.

In group contribution methods groups or building blocks are uniquely designed to generate any possible solvent molecules, and properties of each group can be theoretically calculated, experimentally obtained, or statistically regressed. From a set of groups, all possible combinations of groups can be made to generate molecules. Once molecules are generated, desired properties of the molecules are predicted based on the properties of their groups and tested if they satisfy the pre-specified criteria. If the generated-and-tested molecules have desired properties, they are stored and sorted according to predetermined priorities. This common CAMD approach is called a generation-and-test approach (3,4).

Besides the generation-and-test method, mathematical optimization methods (6,12-14) are usually applied to solve this reverse problem. However, the optimization methods used in these earlier studies are not designed to include uncertainties which can be a major computational bottleneck for the large-scale stochastic combinatorial optimization problems. Here in this paper we are using a combinatorial optimization algorithm specially designed to efficiently handle combinatorial optimization problems under uncertainty (15,16). Figure 2 shows a simple representation of this approach. This approach involves two recursive loops; the outer optimization loop and the inner sampling loop. In this approach the optimizer not only determines the real decision variables like the number of groups in a solvent molecule, group indexes, but also the number of uncertain

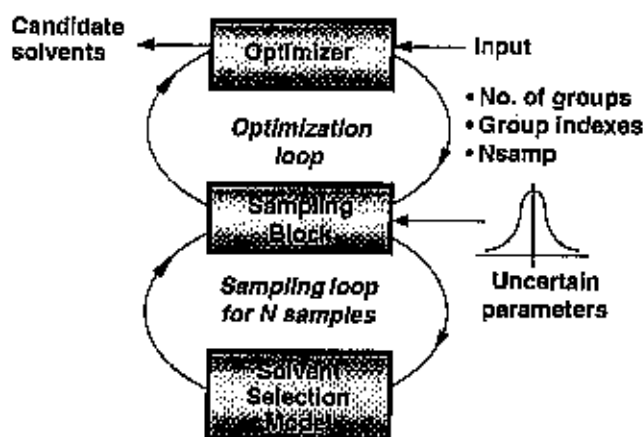


Figure 2. An optimization framework for solvent selection under uncertainty.

samples needed for the inner sampling loop. The inner sampling loop essentially converts the deterministic group contribution model into a stochastic model. Stochastic modeling involves four steps: (a) identifying and specifying key input parameter uncertainties in terms of probabilistic distribution functions, (b) sampling these distributions in an iterative fashion, (c) propagating the effects of these uncertainties through the model, and (d) analyzing the output using statistical techniques.

Total 17 groups and their properties used in this study are shown in Table 2 for estimating solubility parameters. One of the authors (Joback) determined the three solubility parameter terms using the least square method from the literature data (11). In this table, each column consists of group specific solubility parameters and intercept values. The solubility parameter can be estimated by linearly adding group properties. For example, ethyl acetate ( $\text{CH}_3\text{-COO-CH}_2\text{-CH}_3$ ) has three distinctive groups, and its dispersive solubility parameter term is estimated to  $13.38 \text{ MPa}^{1/2}$  ( $2 \times 0.344 + 0.268 - 0.862 + 13.290$ ). Its polar and hydrogen bonding terms are similarly 8.24 and 8.95  $\text{MPa}^{1/2}$ , respectively. The literature values for  $\delta_d$ ,  $\delta_p$ , and  $\delta_h$  are 13.40, 8.60, and 8.90, respectively, and we can see that estimated and literature data are very close.

The set of groups in Table 2 is specially designed for linear or branched hydrocarbons while aromatic, cyclic, and/or halogenated compounds are eliminated due to environmental concerns. As described earlier, one of the features of solubility parameter method is universality: we can use the same

Table 2. Solubility parameters of groups.

Groups	Dispersive	Polar	Hydrogen-bonding
-CH <sub>3</sub>	0.344	-0.591	-0.848
-CH <sub>2</sub> -	0.268	-0.377	-0.595
>CH-	-0.142	-0.801	-1.172
>C<	-1.163	-1.039	-2.496
CH <sub>2</sub> =CH-	-1.163	-1.039	-2.496
CH <sub>2</sub> =C<	-0.243	0.275	-3.542
-CH=CH-	-0.566	-0.034	-0.0776
-CH=C<	-0.695	-0.529	-3.175
>C=C<	-0.823	-1.025	-5.574
-OH	-0.648	5.548	10.630
-O-	-0.638	2.315	1.804
>C=O	-1.145	4.670	4.486
O=CH-	-1.114	5.922	5.256
-COOH	1.068	6.942	11.120
-COO-	-0.862	4.729	4.012
>NH	-1.074	3.875	2.772
-CN	-1.628	6.904	8.317
Intercept	13.290	5.067	7.229

groups for estimating other properties such as boiling point (17), and thus there is no need to have another group set.

#### 4. Uncertainty Identification and Quantification

The Hansen solubility parameters of liquid molecules are estimated by semi-empirical methods, and the three solubility parameter terms for *each group* are regressed using the least square method. Table 3 shows some estimation errors of the Hansen's three-dimensional solubility parameters. For example, ethanol is showing 15% relative error in the solubility parameter mainly due to the error in the  $\delta_b$  term. Estimated solubility parameters of ethyl acetate are quite close to the literature values, which is also shown in Table 1. For diisobutyl ketone, we can find an interesting result. Though each solubility parameter term has large discrepancy, the resulting total solubility parameter is closer to the value reported in the literature. What is important in this study is to quantify uncertainty in each solubility parameter term, not the uncertainty in total solubility parameter, that governs miscibility of two liquids as shown in eq 2.

Table 3. Example calculations of Hansen's solubility parameters.

	Ethanol			Ethyl acetate			Diisobutyl ketone		
	Lit.	Est.	$\Delta(\%)$	Lit.	Est.	$\Delta(\%)$	Lit.	Est.	$\Delta(\%)$
$\delta_d$	12.6	11.7	-7.4	13.4	13.4	-0.1	14.5	13.8	-5.0
$\delta_p$	11.2	9.7	-13.9	8.6	8.2	-4.2	6.8	5.0	-26.2
$\delta_h$	20.0	16.4	-17.9	8.9	9.0	0.6	3.9	5.2	32.0
$\delta$	26.2	22.3	-14.5	18.2	18.1	-0.9	15.5	15.5	-5.8

The group contribution method in this study has 17 groups as shown in Table 2 and each group has three solubility parameter terms. Since the total uncertainty distributions are 51, it is impractical and statistically insignificant to figure out each uncertainty distribution. Instead, the uncertainties of the Hansen solubility parameters ( $\delta_d$ ,  $\delta_p$ , and  $\delta_h$ ) of a liquid molecule are analyzed and quantified in terms of a new parameter called Uncertainty Factor (UF). We define Uncertainty Factor (UF) as the ratio of the literature solubility parameter ( $\delta_{lit}$ ) to the estimated solubility parameter using the group contribution method.

$$UF = \frac{\delta_{lit}}{\delta_{est}} \times 100 (\%) \quad (7)$$

where the UF can be applied to dispersive, polar, or hydrogen bonding terms. Note that the UF of 100 % means that the estimated value is exactly same to the literature value.

In order to elicit the UFs, the estimated solubility parameters of 66 non-cyclic and non-aromatic compounds are compared with the literature values. The probabilistic distributions of the three UFs associated with the three solubility parameters are shown in Figure 3. The UF of the dispersive term ( $\delta_d$ ) is normally distributed with 105.4 % mean and 8.3 % standard deviation. The UFs of  $\delta_p$  and  $\delta_h$  are normally distributed with 121.4 % mean and 128.1 % standard deviation

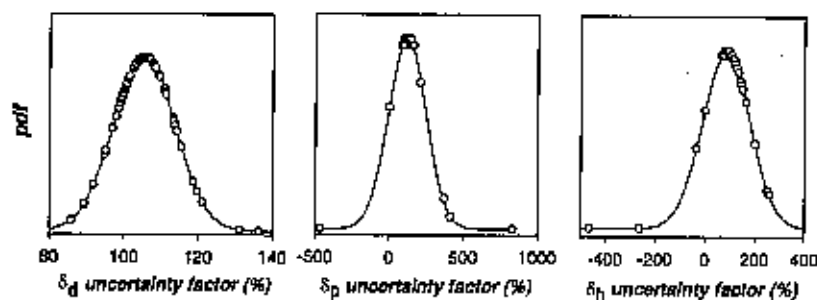


Figure 3. The UFs for dispersive, polar, and hydrogen bonding terms.



and 84.5 % mean and 95.5 % standard deviation, respectively. From this figure, we can see that the effect of uncertainty on  $\delta_h$  is not significant as compared to the effects on the polar term ( $\delta_p$ ) and the hydrogen bonding term ( $\delta_h$ ). Therefore, we can say that UF increases estimated  $\delta_p$  and decreases estimated  $\delta_h$ , and the resulting  $\delta$ ,  $m$ ,  $\beta$ , and  $S_L$  are subject to change.

### Case Study: Solvent Selection for HQAc Extraction

Acetic acid is commonly used as a process solvent or is produced as a byproduct. Therefore, acetic acid is a pollutant as well as a valuable solvent, and it is desirable to minimize the discharge of acetic acid to the environment. To recycle or remove acetic acid from process streams or process units, extraction process is generally applied. For extraction, we can either use high boiling solvents (18,19) or low boiling solvents (4). Ethyl acetate, isoamyl acetate, and isopropyl acetate are widely used in industries to extract acetic acid. This study is mainly focused on finding candidate solvents having low boiling points.

The generation-and-test CAMD approach is usually computer-intensive as this method tries to generate all possible molecular combinations. If the maximum number of groups in a molecule, for example, is 12, then the total number of possible molecular combinations are  $17^{12}$  ( $5.8 \times 10^{14}$ ). In addition, if we consider uncertainties on the solubility parameter terms of groups, this problem becomes computationally expensive.

In order to reduce computational burden and guarantee best candidate solvents, a discrete stochastic optimization method is implemented (15,16) as shown in Figure 2. The optimizer in the upper loop determines the number of groups ( $NG$ ) in a solvent molecule, group indexes ( $N(i)$ ,  $i=1, \dots, NG$ ) that tells which groups are present in a proposed solvent, and the number of uncertain samples ( $N_{\text{sample}}$ ). The information about the distribution functions in terms of 0.1% and 99.9% quantiles is supplied to the inner sampling loop. This sampling loop then uses an efficient sampling method (20) to generate the uncertain samples. Each sample is propagated through the model based on the group contribution method to evaluate the expected values of distribution coefficient, solvent selectivity, and boiling points of each component. This information is transferred to the optimization loop as the objective function and constraints. The optimizer determines whether or not the probabilistic results of the given molecule are optimal.

The main priority (objective function) of solvent selection is distribution coefficient, while the other properties are used as constraints that are summarized in Table 4. The number of groups ( $NG$ ) in a solvent molecule spans

Table 4. Experimental conditions for solvent selection under uncertainty.

Parameter	Bounds
No. of groups in a solvent ( $NG$ )	2 - 12
$\beta$ (solvent selectivity)	$\beta \geq 17.43$
$S_L$ (solvent loss)	$S_L \leq 0.0045$
Boiling point ( $^{\circ}\text{C}$ )	47 - 108

from 2 to 12. The bounds of  $\beta$ ,  $S_L$ , and boiling points are based on the values of ethyl acetate that is one of common solvents for acetic acid extraction.

Table 5 shows the top 20 candidate solvents of the deterministic and stochastic cases. For the deterministic case, the first 10 solvents are alcohols and the remaining solvents are mostly aldehyde. Some of the promising solvents in this case are ethyl alcohol (No. 1), propyl alcohol (No. 3), isopropyl alcohol (No. 5), acetone (No. 12), and methyl ethyl ketone (No. 16). The alcohol function group has the largest hydrogen-bonding term and the second largest

Table 5. Top 20 candidate solvents.

Deterministic case			Stochastic case		
No	Solvent	$m$	No	Solvent	$m$
1	$\text{CH}_3, \text{CH}_2, \text{OH}$	17.2	1	$3\text{CH}_3, \text{CH}, \text{COO}$	161
2	$\text{CH}_3, \text{CH}=\text{CH}, \text{OH}$	12.6	2	$2\text{CH}_3, \text{CH}_2=\text{C}, \text{CO}$	20.1
3	$\text{CH}_3, 2\text{CH}_2, \text{OH}$	9.50	3	$2\text{CH}_3, \text{CH}, \text{OH}$	5.95
4	$\text{CH}_2=\text{CH}, \text{OH}$	9.34	4	$\text{CH}_3, \text{CH}_2, \text{CH}_2=\text{CH}, \text{COO}$	4.91
5	$2\text{CH}_3, \text{CH}, \text{OH}$	6.06	5	$2\text{CH}_3, \text{COO}$	4.26
6	$\text{CH}_2=\text{CH}, 2\text{CH}_2, \text{OH}$	5.83	6	$\text{CH}_3, 2\text{CH}_2, \text{OH}$	3.90
7	$\text{CH}_3, \text{CH}_2=\text{C}, \text{OH}$	5.20	7	$\text{CH}_3, \text{CH}_2, \text{OH}$	3.53
8	$\text{CH}_2=\text{CH}, 2\text{CH}_2, \text{OH}$	3.86	8	$2\text{CH}_3, \text{CO}$	3.51
9	$\text{CH}_3, \text{CH}_2=\text{C}, \text{CH}_2, \text{OH}$	3.41	9	$2\text{CH}_2, \text{CH}_2=\text{CH}, \text{OH}$	3.38
10	$\text{CH}_2, \text{CH}_2=\text{CH}, \text{CH}, \text{OH}$	2.78	10	$\text{CH}_2, \text{CH}_2=\text{CH}, \text{OH}$	3.34
11	$\text{CH}_2, \text{CH}=\text{CH}, \text{CHO}$	2.07	11	$2\text{CH}_3, \text{CH}=\text{C}, \text{CHO}$	2.85
12	$2\text{CH}_3, \text{CO}$	2.06	12	$\text{CH}_2=\text{CH}, \text{OH}$	2.77
13	$\text{CH}_3, 2\text{CH}_2, \text{CHO}$	1.83	13	$\text{CH}_2, \text{CH}=\text{CH}, \text{OH}$	2.58
14	$\text{CH}_3, \text{CH}_2, \text{CH}=\text{CH}, \text{CHO}$	1.50	14	$\text{CH}_2, 2\text{CH}_2, \text{CHO}$	2.40
15	$2\text{CH}_3, \text{CH}, \text{CHO}$	1.47	15	$2\text{CH}_3, \text{CH}_2, \text{CO}$	2.16
16	$2\text{CH}_3, \text{CH}_2, \text{CO}$	1.44	16	$\text{CH}_3, \text{CH}_2=\text{C}, \text{OH}$	2.10
17	$\text{CH}_2=\text{CH}, \text{CH}_2, \text{CHO}$	1.42	17	$2\text{CH}_3, \text{CH}_2, \text{CH}=\text{CH}, \text{O}$	2.09
18	$2\text{CH}_3, \text{COO}$	1.39	18	$2\text{CH}_3, \text{CH}=\text{CH}, \text{CO}$	2.01
19	$\text{CH}_3, 3\text{CH}_2, \text{CHO}$	1.33	19	$2\text{CH}_3, \text{CH}_2=\text{C}, \text{COO}$	1.86
20	$\text{CH}_3, \text{CH}_2=\text{C}, \text{CHO}$	1.22	20	$2\text{CH}_3, \text{CH}_2, \text{CH}, \text{CHO}$	1.85

polar term. This feature of the OH group decreases  $r_{US}$ , resulting in an increase of  $m$ . Similarly the CHO group has the second largest hydrogen-bonding term and the largest polar term which also decrease  $r_{US}$ . Large difference in hydrogen-bonding terms between OH and CHO groups (10.63 vs. 5.26) and small difference in the polar terms (5.55 vs. 5.92) make alcohols preferred solvents for acetic acid extraction.

However, the stochastic case provides a different set of candidate solvents. Only 13 of the solvents generated at the deterministic case are appeared in the list of the stochastic case (See bold numbers in both columns). Some of promising solvents are isopropyl acetate (No. 1), isopropyl alcohol (No. 3), acetone (No. 5), and propyl alcohol (No. 6). Isopropyl acetate that is not listed in the top 20 solvents for the deterministic case, is one of the common industrial solvents for acetic acid extraction and is proved to be highly selective for this extraction purpose. Acetone that is appeared in both cases is reported as the best solvent for this purpose by Joback and Stephanopoulos (4). They also used the solubility parameter method and similar constraints even though their CAMD approach was the generation-and-test method. The combinatorial optimization method used in this study provides more promising solvents than acetone. Ethyl acetate, one of the common industrial solvents, is generated outside the top 20 candidate solvents at both cases since the  $m$  for ethyl acetate (1.04) is relatively low.

Because the mean value of the uncertainty factor of  $\delta_h$  is 84.9%, the contribution of the hydrogen-bonding term decreases in the stochastic case and results in solvents having various functional groups. In addition, the reduced  $\delta_h$  term also decreases the resulting distribution coefficients. If we look at the distribution coefficients of solvents generated in both cases, the expected value of  $m$  under uncertainty is slightly smaller than that of the deterministic case.

Figure 4 shows frequencies of top 40 candidate solvents in both cases. For the deterministic case, as expected, OH and CHO groups are the most common types in the candidate solvents. However, for the stochastic case, other functional groups also have high frequencies, and alkanes and alkenes are in the list of optimal solvents. This means that the stochastic case provide wider range of solvents.

Probabilistic density functions (pdf) of distribution coefficients for both cases are shown in Figure 5. The pdf for the deterministic case is lognormally distributed while the one for the stochastic case is governed by the Weibull distribution with the shape parameter of 1.51 and the scale parameter of 2.58. The Weibull distribution is one of the asymptotic distributions of general extreme value theory; hence, this distribution can approximate the extremely high value of  $m$  of isopropyl acetate under uncertainty. We can also conclude from this figure that the pdf of the stochastic case is narrower and showing

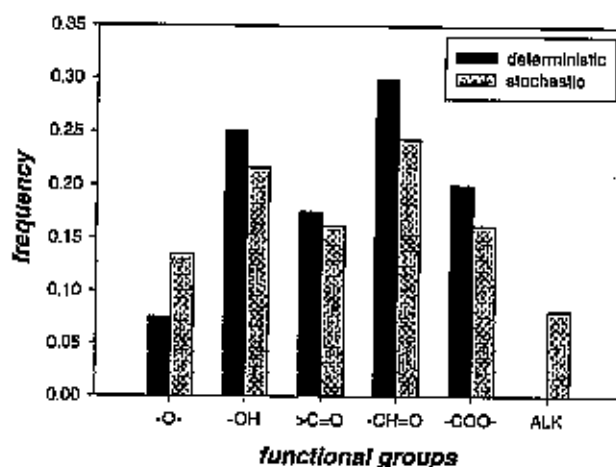


Figure 4. Frequencies of functional groups (ALK means alkanes and alkenes).

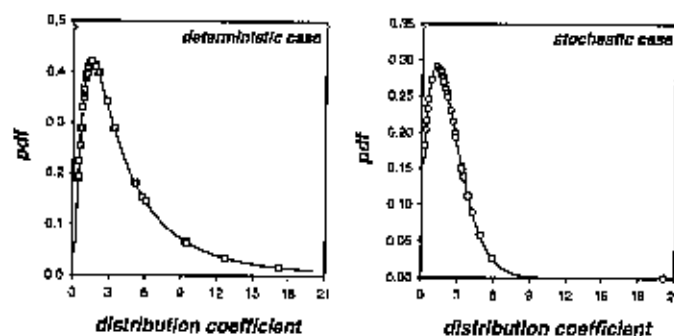


Figure 5. Probabilistic density functions of the distribution coefficient ( $m$ ).

robustness in the solvent selection process. The reduced  $\delta_p$  and the resulting smaller  $m$  can be attributed to this narrow peak of candidate solvents.

For real implementations of CAMD results, solvent selection and its process application (e.g., extraction process) should be simultaneously integrated. In addition, objectives such as process economics, performance, and environmental impacts should be considered. This results in a complex multiobjective optimization programming and represents the current focus for the authors.

### Summary

This paper presents a new combinatorial optimization approach for CAMD under uncertainty. CAMD in this study is based on the group contribution method of the Hansen's three-dimensional solubility parameters. We have used experimental results to quantify uncertainties. A new parameter called uncertainty factor, UF is introduced to quantify uncertainties in the solubility parameter. A real world case study for solvent selection for acetic acid extraction from water is presented to illustrate the approach. For deterministic case, solvents such as ethyl alcohol, propyl alcohol, isopropyl alcohol, acetone, and methyl ethyl ketone are generated. However, under the stochastic case, a different set of solvents such as isopropyl acetate, isopropyl alcohol, acetone, and propyl alcohol are generated. The uncertainty case also identified a commonly used industrial solvent as an important solvent that was not in the top list in the deterministic studies. The analysis showed that the uncertainty results provide wider selection window and robustness in solvent selection. This combinatorial optimization method can also provide more promising solvents than the generation-and-test method.

### Acknowledgement

The authors thank the National Science Foundation for their funding this research (CTS-9729074).

### References

1. The Montreal Protocol on Substances that Deplete the Ozone Layer. 1987.
2. Zhao, R.; Cabezas, H. *Ind. Eng. Chem. Res.* 1998, 37, 3268.
3. Gani, R.; Brignole, E.A. *Fluid Phase Equilibria* 1983, 13, 331.
4. Joback, K.G.; Stephanopoulos, G. *Proc. FOCCAPD '89*, Snowmass, CO, 1989, pp 363.
5. Buxton, A.; Livingston, A.G.; Pistikopoulos, E.N. *AIChE Journal*, 1999, 45, 817.
6. Vaidyanathan, R.; El-Halwagi, M. *Ind. Eng. Chem. Res.* 1996, 35, 627.
7. Constantinou, L.; Jakslund, C.; Bagherpour, K.; Gani, R.; Bogle, I.D.L. *AIChE Symposium Series*, AIChE: New York, 1994, Vol. 90, No. 303; pp 105.
8. Panayiotou, C. *Fluid Phase Equilibria*, 1997, 131, 21.
9. Hildebrand, J.; Scott, R.L. *Regular Solutions*, Prentice-Hall; Englewood Cliffs, NJ, 1962.

10. Hansen, C.M. *Ind. Eng. Chem. Prod. Res. Dev.* **1969**, *8*, 2.
11. Barton, A.F.M. *CRC Handbook of Solubility Parameters and Other Cohesion Parameters*, CRC Press: Boca Raton, FL, 1983, pp 94.
12. Odele, O.; Macchietto, S. *Fluid Phase Equilibria*, **1993**, *82*, 47.
13. Maranas, C.D. *Ind. Eng. Chem. Res.* **1996**, *35*, 3403.
14. Venkatasubramanian, V.; Chan, K.; Caruthers, J.M. *Comp. Chem. Eng.* **1994**, *18*, 833.
15. Chaudhuri, P.; Diwekar, U.M. *AIChE J.* **1996**, *42*, 742.
16. Chaudhuri, P.; Diwekar, U.M. *AIChE J.* **1999**, *45*, 1671.
17. Joback, K.G.; Reid, R.C. *Chem. Eng. Commun.* **1987**, *57*, 233.
18. Pretel, E.J.; Lopez, P.A.; Bottini, S.B.; Brignole, E.A. *AIChE J.* **1994**, *40*, 1349.
19. Hostrup, M.; Harper, P.M.; Gani, R. *Comp. Chem. Eng.* **1999**, *23*, 1395.
20. Kalagnanam, J.R.; Diwekar, U.M. *Technometrics* **1997**, *39*, 308.