A systematic molecular design framework for an environmentally benign solvent recovery process

Jecksin Ooi[1], *Denny K. S.* Ng[1], and *Nishanth* Chemmangattuvalappil[1,*]

[1]Department of Chemical and Environmental Engineering/Centre of Excellence for Green Technologies, The University of Nottingham Malaysia Campus, Broga Road, 43500 Semenyih, Selangor, Malaysia

Abstract. Computer Aided Molecular Design (CAMD) techniques have been extensively applied to design solvents for different applications. Most of the CAMD problems only aim at generating solvents that meet the predefined functionality. Nevertheless, it is important to consider the effect of solvent on the safety, health and environmental impacts during the recovery process. This paper presents a single stage CAMD framework that simultaneously quantifies the environmental impact of the solvent recovery process. The environmental impact of the process can be estimated through IChemE Sustainability Metrics. Besides, molecular properties that have an impact on the quantitative evaluation of the environmental impact of solvent recovery process are included in this framework. Weighted sum method coupled with Fuzzy Analytic Hierarchy Process (FAHP) weighting approach is employed to solve the multi-objective molecular design framework. A case study on solvent design for residual oil extraction from palm pressed fiber is presented to illustrate the proposed framework. In this work, only energy balance around multistage evaporator is incorporated into CAMD formulation as energy required to recover the solvent contributes to the largest portion of the whole process. The results show that the designed solvents simultaneously possess target functionalities and reduce the environmental impact of solvent recovery process.

1 Introduction

Solvents are heavily applied throughout the chemical industries for various purposes, such as being a medium for chemical production, separation processes and chemical reactions. Most of these solvents are organic compounds which are hazardous and lead to adverse impacts on the environment when being released to the atmosphere. Hence, separation techniques are important for solvent recovery from the waste stream to minimize pollution and improve economic performance **[1]**. Solvent recovery process usually requires excessive amount of energy which will results in environmental issues such as global warming. Owing to increasingly stringent environmental regulations, most chemical industries actively look for non-hazardous and environmental benign solvents. It is of vital importance to choose a suitable solvent because the process performance is highly affected by the solvent applied in it **[2]**.

According to Capello et. al **[3]**, various environmental effects have been found to be interrelated with the use of solvents in industries. For instance, solvent incineration often causes air pollution whereas solvent recovery process needs large amount of energy. Thus, this has shown that the performance and environmental impact of a solvent recovery process are heavily influenced by the physical feature of a solvent. In another word, physicochemical properties of a solvent applied in a unit operation dictates the difficulty level of a chemical process. Higher degree of difficulty of a chemical process often possesses higher risk of causing SHE issues. To date, many industries only focus on solvent screening when comes to solvent selection. However, this practice is time-consuming because exhaustive search must be carried out from a large database to identify a potential solvent. Hence, a systematic methodology to determine solvents which excel in targeted properties and reduce environmental impact of its recovery process is needed.

Computer Aided Molecular Design (CAMD) technique emerges as a promising systematic route to determine solvents that have to be tested through experimental works. This is mainly because through CAMD, larger set of solvent properties can be evaluated rapidly and cost-effectively through computational methodology. CAMD technique is a reverse engineering method, aims to generate molecules with a set of molecular building blocks and predefined target

* Corresponding author: Nishanth.C@nottingham.edu.my

properties **[4]**. Throughout the years, CAMD techniques have been widely employed to generate solvents which act as process liquids or extracting agents in various process, solvent-based and/ or pharmaceutical industries **[5]**. Nonetheless, no work is found on reporting a framework that can simultaneously generate solvent that achieves its functionalities and reduce the environmental impact of its recovery process. To bridge the research gap, a novel methodology is proposed to perform quantitative assessment on the effect of designed solvent on environmental effect from molecular level to process stage. This work explores the application of IChemE Sustainability Metrics to calculate the total environmental burden (EB) of solvent recovery process during molecular design. Via such method, a solvent that fulfils favorable functionalities and improve the overall environmental performance of the recovery process can be generated.

2 Methodology

This work presents a novel CAMD methodology for the design of molecule which simultaneously exerts targeted properties and minimize SHE effects of its recovery process. The quantitative assessment of environmental effect of a solvent recovery process is integrated into the CAMD framework. Fuzzy Analytic Hierarchy Process (FAHP) approach is employed to systematically determine the weighting factor of each property. FAHP is an extension of Analytic Hierarchy Process (AHP), which was first developed by Saaty **[6]**. FAHP approach employs fuzzy set theory instead of discrete numerical values to capture the uncertainty and vagueness of decision makers' opinion **[7].** Besides, the degree of confidence of decision maker can be quantified through triangular fuzzy number when eliciting judgements **[8]**. This approach can generate a molecule with good safety and environmental feature that attains its functionalities and improve the environmental characteristic of its solvent recovery process. The developed methodology has been summarized in the following steps.

2.1 Problem formulation

In this stage, the design objective is first specified for CAMD problem. This is normally done by determining the product needs based on customers' preferences or from the operating condition of industrial processes. In addition to physicochemical properties, safety, health and environmental (SHE) properties are considered to design a molecule with preferable functionalities and fulfils SHE criteria. Besides, from a process perspective, the generated molecule should be able to minimize adverse effect on the environment. The quantitative evaluation of environmental impact of a solvent recovery process is performed by using the steps shown in section 2. The target properties that are selected as target properties are either estimated through property prediction models or empirical correlations. In this work, first order group contribution method (GCM) equation, which is shown in **Eq. 1** is chosen as property prediction model. The function $f(p)$ of the target property p is shown by the left- hand side of Eq. (1) whereas C_i represents the contribution of the group of type-i that occurs N_i times.

$$f(p) = \sum_i N_i C_i \tag{1}$$

2.2 Quantitative assessment of environmental impact of a solvent recovery process

The first step involves identifying the critical parameters required for the calculation of heat balance for a specific process. Next, based on the unit operations of the process, potential energy sources such as electricity, steam or natural gas is determined. By performing this step, energy required for a recovery process can be converted into greenhouse gases (GHG) emissions such as methane (CH_4), carbon monoxide (CO), carbon dioxide (CO_2), nitrous oxide (N_2O), nitrogen oxide (NO_x), sulfur oxide (SO_x) and volatile organic compound (VOC) based on the emission factors extracted from LCA Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET) model 2016 database. **Table 1** shows the emission factors for converting energy consumption into GHG emissions.

Table 1. Emission factors for converting energy consumption into GHG emissions.

GHG (g)	Natural Gas (1 tonne)	Steam (1 MJ)	Electricity (1 kWh)
CO	540	0.0677	0.0696
CO_2	219,570	78.360	120
CH_4	4620	0.3400	0.46
NOx	750	0.0965	0.0870
SO_X	510	0.0142	0.0216
N_2O	1.54	0.0021	0.0017
VOC	290	0.0153	0.0164
References	**[9]**	**[9]**	**[10]**

IChemE Sustainability Metrics, developed by the Institution of Chemical Engineers (IChemE) are applied to evaluate the environmental performance of recovery process. The environmental performance of recovery process is evaluated by calculating EB for various environmental effects such as acidification potential, eutrophication potential, global warming and etc. The EB equivalent for different types of pollution is referred from IChemE sustainability metrics **[11]**. **Table 2** shows the EB for each environmental effect, which is given by comparing the effect of a pollutant to that of a standard substance. EB of a recovery process can be expressed mathematically in **Eq. 2**:

$$EB = \sum M_j \times PR_{b,j} \tag{2}$$

where EB is the total environmental burden; M_j is the mass of pollutant j emitted; $PR_{b,j}$ denotes the potency factor of pollutant j for the environmental impact, b.

The potency factor of specific pollutant j is extracted from IChemE sustainability metrics **[11],** which is shown in **Table 3**.

Table 2. EB equivalent for various types of pollution **[11]**

Impacts	Pollution	EB equivalent
Atmospheric	Atmospheric acidification	SO_2
	Global warming potential	CO_2
	Human carcinogenic effects	Benzene
	Stratospheric ozone depletion	CFC-11
	Photochemical ozone formation potential	Ethylene
Aquatic	Aquatic acidification	Released H^+ ions
	Aquatic oxygen demand	Oxygen
	Ecotoxicity to aquatic life	Copper
		Formaldehyde
	Eutrophication	Phosphate

Table 3. Potency factor of each pollutant **[11]**

Pollutant	Global warming potential	Photo-chemical ozone formation potential	Atmospheric Acidification	Eutrophi-cation
CO_2	1	-	-	-
VOC	11	-	-	-
CO	3	0.027	-	-
NO_x	40	0.028	0.17	0.13
SO_x	0	-	-	-
CH_4	21	-	-	-
N_2O	310	-	-	-

2.3 Molecular design and optimization model

In molecular design stage, suitable molecular building blocks are chosen based on the nature of the CAMD problem. Structural constraints are then imposed to ensure a structurally feasible molecule is formed. In this work, only simple structured acyclic compounds are considered.

CAMD problem is formulated as a multi-objective optimization problem since there are more than two targeted properties being chosen as design objectives. By referring to the previous works done by Ooi et al. **[12-13]**, weighted sum method can be applied to solve multi-objective CAMD problem. Weighted sum method then transforms multiple objectives into an aggregated scalar objective function. Owing to that the objective functions belong to various categories such as SHE and physicochemical properties, these properties are represented by different measurement scales and units. Hence, normalization step is performed to bring them to the same scale. The step is then followed by assigning each objective function with the weightage identified from FAHP approach, and then add up all the contributors to obtain the overall objective function. The overall objective function is expressed by **Eq. 3**:

$$F^{weighted\ sum} = w_1\lambda_{p1} + w_2\lambda_{p2} + \ldots + w_m\lambda_{pm} \quad (3)$$

where $F^{weighted\ sum}$ is the overall objective function and w_m is the weighting factor for each normalized target property operator λ_{pm}. The design objective of this work is to maximize $F^{weighted\ sum}$. In order to generate feasible solution, integer cuts are applied.

3 Case study

3.1. Problem formulation

Palm Pressed Fiber (PPF) is a biomass by-product formed after the extraction of crude palm oil (CPO) from fresh fruit bunches (FFB). This case study aims at designing alternate solvents to replace hexane for extracting residual oil from PPF. The designed solvents are expected to exert favorable features from both molecular and process aspects. The designed solvents should have low EB for their recovery process in order to minimize the impact on the environment. The solvents are assumed to be recovered through multistage evaporator. Since the core energy requirement of the process comes from multistage evaporator, energy balance around multistage evaporator is integrated into CAMD formulation. The energy balance around multistage evaporator is shown in **Eq. 4**:

$$M_{steam}\lambda_{steam} = M_s C_{ps}(T_{evp} - T_{feed}) + M_s H_{vs} + M_{oil} C_{poil}(T_{evp} - T_{feed}) \quad (4)$$

where M_{steam} is the mass of steam needed (kg); λ_{steam} is the latent heat of steam (kJ/kg); M_s is the mass of designed solvent fed into evaporator (kg); C_{ps} is the average heat capacity of solvent (kJ/kg °C); T_{evp} and T_{feed} represent evaporator and feed temperature respectively (°C); H_{vs} is heat of vaporization of the solvent (kJ/kg); M_{oil} and C_{poil} is the mass (kg) and specific heat capacity (kJ/kg °C) of residual oil. The environmental effect of solvent recovery process can then be quantified and integrated into CAMD problem by using **Eq. 1** and **Eq. 4** together with information from **Table 1**.

Moreover, to ascertain that solvents will exert good functionalities, the solvents should have low boiling point (T_b) to reduce the degradation rate of carotene, low viscosity (μ) and surface tension (σ) to promote solvent diffusivity as well as small difference of Hildebrand solubility parameter (δ) between solvent and carotene ($\delta_{carotene}$) to ensure that both carotene and triglycerides (TAGs) is highly soluble in the solvents. Besides, to lessen SHE risks, the solvent itself should have the smallest total penalty score for safety and health aspects (I_{SHI}), small soil sorption coefficient (log K_{oc}), small bioconcentration factor (BCF), low potential in causing photochemical oxidation (PCO) and aquatic toxicity

(high LC_{50} toxicity). I_{SHI} of a molecule can be calculated using **Eq. 5**.

$$I_{SHI,w} = w_{FL} I_{FL} + w_{EX} I_{EX} + w_{EL} I_{EL} + w_{AH} I_{AH} \quad (5)$$

All the properties that are selected as target properties and to calculate EB and I_{SHI} can be estimated through property prediction models. The selected molecular groups are C, CH, CH_2, CH_3, OH, COOH, CHO, CH-O, CH_3CO, CH_3O, CH_2O, CH_3COO, NH_2 and CH_2=CH. After determining weightage for each objective function, the overall objective function can be defined as Eq. (6). The design objective of this CAMD problem is to maximize the value of $F^{weighted\ sum}$.

$$F^{weighted\ sum} = 0.0809\lambda_{T_b} + 0.0294\lambda_{\sigma} + 0.0294\lambda_{\mu} + 0.0156\lambda_{R_{carotene}} + 0.0077\lambda_{LC_{50}} + 0.0278\lambda_{PCO} + 0.0034\lambda_{\log K_{oc}} + 0.0034\lambda_{BCF} + 0.1302\lambda_{I_{FL}} + 0.1302\lambda_{I_{EX}} + 0.0211\lambda_{I_{EL}} + 0.0211\lambda_{I_{AH}} + 0.5\lambda_{EB} \quad (6)$$

3.2 Results and discussion

In this work, ten acyclic solvents with the highest $F^{weighted\ sum}$ value are generated. **Fig. 1** shows the molecular structures of the top three generated acyclic solvents. The properties of the top three solvents are then compared with that of hexane, which is reported in **Table 4**. From **Table 4**, it shows that T_b of the top three solvents are comparable or lower than that of hexane. Besides, these three solvents have significantly smaller $\delta_{carotene}$ values than hexane which implies that carotene will be highly soluble in these generated solvents. Furthermore, solvent 1 and solvent 2 exert lower total EB compared to that of hexane as the mentioned solvents consume relatively lower amount of energy during solvent recovery process. The designed solvents will be safer than hexane because they have higher flash point. In addition, these solvents will be less likely in causing aquatic toxicity due to their lower – log LC_{50} value. Comparison between the generated solutions and the existing solvents in literature has been made. It depicts that solvent 3 which is ethyl acetate provides a better carotene yield (42.61 mg/kg) compared to that of hexane (34.45 mg/kg) **[14]**. Nevertheless, this generated ranking is not absolute. These identified potential molecules will still need to be sent for further verification through experiments in a later stage.

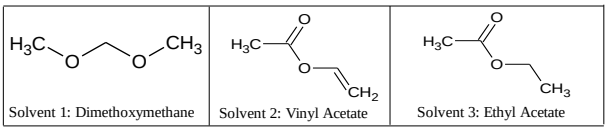

Fig. 1. Molecular structures of the top three acyclic solvents.

Table 4. Properties of hexane and top three solvents

	Hexane	Solvent 1	Solvent 2	Solvent 3
Boiling Point (°C)	68.7	48.5	66.9	68.7
Surface Tension (mN/m)	17.9	20.2	24.6	23.5
Viscosity (cp)	0.30	0.31	0.36	0.43
$\delta_{carotene}$	2.9	0.31	0.49	0.92
Flash Point (°C)	-30.9	-6.99	3.36	4.5
S (vol%)	7.01	14.74	13.4	12.3
PEL(ppm)	50	1.16	7.41	194.5
-log LC_{50}	4.54	-2.95	-0.72	-1.37
LD_{50} (mg/kg)	28000	711.06	840.82	1236.33
PCO	0.431	0.451	0.744	0.279
BCF	51.36	3.98	14.79	14.43
log k_{oc}	3.62	-0.284	-0.062	0.175
EB	6.931	6.015	6.785	7.006
I_{SHI}	0.563	0.605	0.605	0.563

4 Conclusion

This paper presents a novel CAMD methodology that design solvents with optimal targeted properties and being recovered without causing excessive environmental burden. This approach utilizes IChemE Sustainability Metrics to estimate the total environmental burden of solvent recovery process. To illustrate the proposed method, a case study on the solvent design for the extraction of residual oil from PPF is solved. In this case study, properties from both molecular and process aspects are selected as objective functions. Weighted sum method together with FAHP weighting approach is then applied to solve multi-objective CAMD problem. The results dictate that the solvents possess a good balance between process and molecular properties. The energy consumption as well as environmental burden of solvent recovery process can be reduced by using the generated solvents. Future work can be carried out by extending the CAMD methodology to consider business aspects.

This research was financially supported by Ministry of Higher Education, Malaysia through the LRGS Grant (LRGS/2013/UKM-UNMC/PT/05).

References

[1] Ramzan, N., Degenkolbe, S. and Witt, W. (2008) 'Evaluating and improving environmental performance of HC's recovery system: A case study of distillation unit', *Chemical Engineering Journal*, 140(1–3), pp. 201–213.

[2] Papadopoulos, A. I. and Linke, P. (2009) 'Integrated solvent and process selection for separation and reactive separation systems', *Chemical Engineering and Processing: Process Intensification*, 48(5), pp. 1047–1060.

[3] Capello, C., Hellweg, S. and Badertscher, B. (2007) 'Environmental Assessment of Waste-Solvent Treatment Options', *Journal of Industrial Ecology*, 11(4), pp. 26–38.

[4] Harper, P. M. and Gani, R. (2000) 'A multi-step and multi-level approach for computer aided molecular design', *Computers & Chemical Engineering*, 24(2), pp. 677–683.

[5] Gebreslassie, B. H. and Diwekar, U. M. (2015) 'Efficient ant colony optimization for computer aided molecular design: Case study solvent selection problem', *Computers & Chemical Engineering*, 78, pp. 1–9.

[6] Clark, C. E., Han, J., Burnham, A. J., Dunn, J. B. and Wang, M. Q. (2011) 'Life-Cycle Analysis of Shale Gas and Natural Gas'. Argonne, IL USA, p. 52

[7] Cai, H., Wang, M., Elgowainy, A. and Han, J. (2012) *Updated Greenhouse Gas and Criteria Air Pollutant Emission Factors and Their Probability Distribution Functions for Electric Generating Units*.

[8] Saaty, T.L., (1977) 'A scaling method for priorities in hierarchical structures', *Journal of Mathematical Psychology*, pp. 234 – 281.

[9] Zadeh, L.A., (1965) 'Fuzzy sets', *Information and Control*, 8, pp 338–353.

[10] Tan, R.R., Aviso, K.B., Huelgas, A.P., Promentilla, M.A.B., (2014) 'Fuzzy AHP approach to selection problems in process engineering involving quantitative and qualitative aspects', *Process Safety and Environmental Protection*, 92, pp 467–475.

[11] IChemE (2002) *The Sustainability Metrics*. Available online at http://nbis.org/nbisresources/metrics/triple_bottom_line_indicators_process_industries.pdf

[12] Ooi, J., Promentilla, M. A. B., Tan, R. R., Ng, D. K. S. and Chemmangattuvalappil, N. G. (2017) 'A systematic methodology for multi-objective molecular design via Analytic Hierarchy Process', *Process Safety and Environmental Protection*, 111(Supplement C), pp. 663–677.

[13] Ooi, J., Promentilla, M. A. B., Tan, R. R., Ng, D. K. S. and Chemmangattuvalappil, N. G. (2018) 'Integration of Fuzzy Analytic Hierarchy Process into multi-objective Computer Aided Molecular Design', *Computers & Chemical Engineering*, 109 (Supplement C), pp. 191–202.

[14] Strati, I. F. and Oreopoulou, V. (2011) 'Effect of extraction parameters on the carotenoid recovery from tomato waste', *International Journal of Food Science & Technology*. Blackwell Publishing Ltd, 46(1), pp. 23–29.