MODELLING OF SUPERCRITICAL FLUID EXTRACTION USING DYNAMIC GENETIC ALGORITHM BASED OPTIMISATION

Simon X. Yang [*,**] **Jin Zeng** [**] **Guoyin Wang** [*]

[*] *Institute of Computer Science and Technology, Chongqing University of Posts and Telecommunications, China*
[**] *School of Engineering, University of Guelph, Canada*

Abstract: In food and chemical industry, supercritical fluid extraction is an environmental-friendly separation technique, which plays an important role in biomaterial processing. In this paper, a dynamic optimisation model for supercritical fluid extraction is proposed by combining a transformation based genetic algorithm and the Peng-Robinson equation of state. The proposed dynamic model for the relationship between pressure and yield of biomaterial can recognise the change of temperature, and is able to adapt to a new solution by finding a near-optimal k12 parameter without restarting the system as other models. The effectiveness of the proposed approaches is demonstrated by simulation and comparison studies.

Keywords: Dynamic modelling, Supercritical fluid extraction, Genetic algorithm

1. INTRODUCTION

Supercritical fluid extraction (SFE) is a novel separation technique in food science and chemical industry, which exploits the solvent properties of fluids near the critical point, extracts the material using supercritical fluid (Kaihara *et al.*, 2002; de Lucas *et al.*, 2002; Oliveira *et al.*, 2002). During the separation process, the yield rate and the solubility of biomaterials can be controlled by adjusting some specific parameters such as pressure, temperature, moisture contents, etc. However, under the experimental conditions, it is extremely difficult and costly to obtain the high-pressure phase-equilibrium data for supercritical fluid. Thus it is desirable to find a reliable model to extend and complement the experimentally obtained data at high pressure for a SFE-solute system. In the past years, several models for supercritical fluid extraction have been proposed to describe the extraction rate by various mathematical approaches (Clifford, 1998; Ferreria *et al.*, 1993; Goto *et al.*, 1993; Naik *et al.*, 1989; Sovava, 1994). However, there are two major problems in developing of a proper mathematical model, one problem is the lack of understanding of the dense fluid state, the other is that the interaction of molecules differs significantly in molecular size, shape and polarity.

In order to develop a reliable model for SFE, recently years neural network based and evolutionary algorithm based techniques have been applied to this area. **?** first proposed models for SFE using the multi-layer perceptron neural network and the radial basis function network. Both models can produce more accurate and reliable predictions of the relationship between pressure and yield of biomaterials, without any tedious selection of model parameter by trial and error as the conventional models do. Then, Li and Yang's (2003) proposed another hybrid model for SFE using genetic algorithms and Peng-Robinson equation of state (PR). In this model, the genetic algorithm is

used to select the near-optimal binary interactive parameter to fit the experimental data.

Although the above models using neuro-based and evolutionary-based techniques can solve the problems associated with the mathematical models, they still have shortcomings. Due to the properties of binary interaction parameter k_{12}, those models only suit for a certain temperature, not suit for the whole dynamic process. Once the temperature changed during the extraction process, those models unable to adjust themselves to fit the new environment, they need to restart the system, modify some parameters in order to get a new model to fit the extraction process in the changed temperature. In this case, the standard approach is to regard each temperature change as the arrival of new optimisation problem that has to be solved from scratch. This approach however if often impractical, because the solution to the new problem should not differ too much from the solution of the old one. So, develop a dynamic model for the extraction process is very useful.

Traditional genetic algorithms (GA) are not suitable to solve problems in a dynamic environment, because the population converges to optimum very quickly and lose of the diversity, when changes occur, it is very difficult to re-adapt the solutions to the new conditions. In this study, a dynamic model that combines the transformation base GA (TGA) and the Peng-Robinson equation of state is proposed to model the dynamic process of supercritical fluid extraction. In our simulation study, this dynamic model can always detect the change of temperature and help the system to select the near-optimal binary interaction parameter k_{12} for the Peng-Robinson equation of state, it can provide more accurate and reliable predictions of the extraction process.

2. THE MODEL ALGORITHMS

In this section, a typical Peng-Robinson equation of state and the transformation based GA which we used in this study are introduced, then the proposed Hybrid Dynamic TGA-PR model is presented.

2.1 A Typical Peng-Robinson Model

In the past decades, a large number of complex and realistic equations of state have been proposed. The Peng-Robinson equation of state is the most commonly used method for treating solubility in SCF. This equation has been used to predict phase behaviour for solutes with a wide range of volatility. In this method, it is assumed that the solute at equilibrium with the saturated solution is a solid, which contains a negligible amount of the supercritical fluid substance. The partial molar volume of the solid solute in the system at all the considered pressures and temperatures is equal to its molar volume V at the atmospheric pressure and the temperature of 298K. The Peng-Robinson equation of state is given as

$$p = \frac{RT}{V - b_1} - \frac{a_{11}(T)}{V^2 + 2Vb_1 - b_1^2}, \tag{1}$$

where R is the gas constant, T is the absolute temperature, V is the molar volume of the pure solvent, a is the parameter describing attractive interactions between molecules, and b is the parameter describing volume exclusion and repulsive interactions. Subscript 1 represents the solvent and 2 represents the solute. Parameters a and b are determined from the critical properties of the components according to the following equations,

$$a = \frac{0.45724R^2T_c^2}{p_c}\Big[1+f(w)(1-\sqrt{T_r})\Big]^2, \tag{2}$$

$$f(w) = 0.37464 + 1.54226w - 0.26992w^2, \tag{3}$$

$$T_r = \frac{T}{T_c}, \tag{4}$$

$$b = \frac{0.07780RT_c}{p_c}, \tag{5}$$

where T_c and p_c are the critical temperature and critical pressure, respectively, and w is the acentric factor. Thus the pure component parameters, a_i and b_i, can be calculated as

$$a_{11} = \frac{0.45724R^2T_{c,1}^2}{p_{c,1}}\Big[1+f(w_1)(1-\sqrt{T_{r,1}})\Big]^2, \tag{6}$$

$$b_1 = \frac{0.07780RT_{c,1}}{p_{c,1}}, \tag{7}$$

$$b_2 = \frac{0.07780RT_{c,2}}{p_{c,2}}. \tag{8}$$

In order to calculate the properties of a mixture, the parameter k_{12} that describes the mixture must be introduced. Parameter k_{12} is called a binary interaction parameter and is used to account for the non-ideality of attraction between 1 and 2. In general, k_{12} should have a small positive value between 0 and 1. The binary interaction parameter is often adjusted to make the equation of state fit the experimental data. Therefore it can be derived as

$$a_{12} = \frac{0.45724(1-k_{12})R^2T_{c,1}T_{c,2}}{\sqrt{p_{c,1}p_{c,2}}} \Big[1 + f(w_1)(1-\sqrt{T_{r,1}})\Big] \Big[1 + f(w_2)(1-\sqrt{T_{r,2}})\Big]. \tag{9}$$

The Peng-Robinson equations for the mole fraction, x_2, at saturation of a solute of low volatility in SCF can be written as

$$\ln x_2 = \ln \frac{p_v(T)}{p} + \frac{pV_m}{RT} - \ln \phi_2, \qquad (10)$$

where $p_v(T)$ is the vapour pressure of the solute, and V_m is the volume of the pure solute, and ϕ_2 is the fugacity coefficient, which can be calculated as

$$\ln \phi_2 = -\frac{a_{11}}{2\sqrt{2}RTb_1}\left(\frac{2a_{12}}{a_{11}} - \frac{b_2}{b_1}\right) \ln \frac{V + (1+\sqrt{2})b_1}{V + (1-\sqrt{2})b_1} + \ln \frac{RT}{p(V-b_1)} + \frac{b_2}{b_1}\left(\frac{pV}{RT} - 1\right). \qquad (11)$$

These equations allow the calculation of solubility at a given temperature and solvent molar volume, given the vapour pressure and molar volume of the state, the critical parameters and acentric factors of both components and the binary interaction parameter.

2.2 Overview of Transformation Based Genetic Algorithms

The idea of Transformation based Genetic Algorithms is come from the biological transformation in the nature. Such as a virus infect a cell, first use its own DNA segment to replace some of the host's DNA segment, then after recombination and integration to take over the whole host. The TGA is first proposed by Simöes and Costa (Simöes and Costa, 2001), this algorithms is suitable for solving dynamic optimisation problems.

The TGA is different with the traditional GA, it introduced a new operator "transformation" which different with the traditional operator "crossover". Firstly, it random generate an initial population and an initial gene segment pool, the genes length in the gene segment pool are different. Then, in each generation, individuals are selected and be modified by the gene segments in the gene segments pool, this process called "transformation", which replaced the traditional operator "crossover". After transformation and mutation, a new population is created. Then using this new population to replace the old one, and also using the old population to create part of the new gene segment pool, the remaining part are random generated. Then go on to the next generation until the maximum generation or the stop criteria reached.

For the transformation process, a gene segment is selected from the segment pool with a fixed probability, and a start point of transformation in the selected individual is also selected at random. The gene segment is replacing the gene in the individual after the start point. Here, the chromosome is treated as a circle to keep the chromosome length.

TGA has some advantages than traditional GA. Traditional GA are often used to solve stationary problems, it has some limitations to deal with the dynamic problems. When the environment changes, the traditional GA can not adapt the changes because the lost of diversity in the population. In this point, TGA is always keep the population's diversity, it can recognised the change of environment and was able to adapt to the new solution without restart the system as the traditional GA does.

2.3 The Dynamic Modelling for the Supercritical Fluid Extraction

As we mentioned in the previous section, although the neural network and the traditional GA can successfully used in modelling the SFE, select the optimal binary interaction parameter k_{12} for Peng-Robinson equation to fit the experimental data. However, those models only suit for a certain temperature, not suit for the whole dynamic process. That means, during the extraction process, if the reactive temperature changes, we need to modify some parameters in the those models in order to get a new model to fit the extraction process for the changed temperature. This approach is impractical, so, here we use transformation base GA to realise the dynamic modelling of SFE. The schematic diagram of the proposed hybrid dynamic model is shown in Fig. 1.

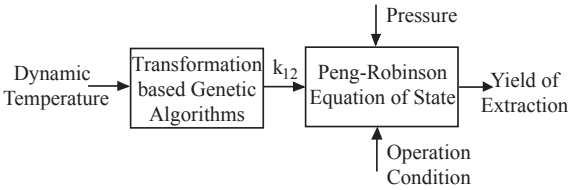

Fig. 1. Schematic diagram of the proposed hybrid dynamic model for SFE.

The hybrid dynamic model correlated by TGA is summarised as follows:

1. Generate the initial population of individuals and initial gene segment pool;
2. Encode k_{12} to get chromosomes' evaluation value for initial generation;
3. Encode k_{12} to get evaluation value for one generation;

   3.1 Calculate yield rate by using Peng-Robinson equation with k_{12}, increase the

pressure, repeat to get all values for the pressure range;

3.2 Calculate the mean square error and correlation, then get the corresponding evaluation value;

3.3 Evaluate the current generation: (1) if the fitness value satisfied, jump to step 11; (2) not satisfied and not all the individuals evaluated, repeat step 3; and (3) not satisfied and all the individuals evaluated, continue to step 4;

4. Select individuals from the population to a mating pool, prepare for transformation and random select a transform start point;
5. Select one gene segment from the gene segment pool;
6. Do the transformation, use the gene segment to replace the gene in the individual after the start point;
7. Do the mutation as the traditional GA does and get the new offspring;
8. Generate new population

   8.1 Population size not reached, back to step 4;

   8.2 Population size reached, continue to step 9;
9. Update the gene segment pool

   9.1 Select individuals from the old population to create part of the new gene segment pool;

   9.2 Random select the gene segment length and the start point in individuals;

   9.3 Create new gene segment;

   9.4 Remaining part of gene segment pool are created by random;
10. Replace the old population with new population, back to step 3;
11. Store the k_{12} selected by TGA, calculate the yield rate using Peng-Robinson equation.

In this study, the chromosome selected is a 12 bits binary string, which represent the k_{12} in decimal value between 0 to 1. The population size and the gene segment pool size are defined of 30 individuals. The individuals are selected to be transformed by using the tournament selection mechanism with a fixed probability (75%). Part of the gene segment pool (70%) is updated every generation, using genetic information for the individuals of the old population, also selected by using the tournament selection method with a fixed probability (75%), the rest part (30%) of the gene segment pool are updated by random. The mutation process selects a random position of the individuals to negates the bit value with fixed probability (5%).

For the fitness function, in this study, the least mean square error is selected to evaluate the TGA. The fitness function defined as

Table 1. Solubility of benzoic acid in supercritical CO_2 extraction, at $35°C$, $45°C$, $55°C$ and $70°C$, respectively (p: pressure, bar; γ_y: yield rate, mole fraction, $\times 10^{-3}$).

$35°C$		$45°C$		$55°C$		$70°C$	
p	γ_y	p	γ_y	p	γ_y	p	γ_y
101.0	0.80	101.0	0.26	101.0	0.14	101.0	0.122
113.0	1.03	105.0	0.493	105.0	0.22	111.0	0.222
120.0	1.38	113.0	0.561	111.0	0.30	126.0	0.55
151.0	1.74	120.0	1.15	120.0	0.49	151.2	1.54
160.0	2.37	151.0	1.98	126.0	0.75	200.9	5.00
240.0	3.10	160.0	2.38	151.2	1.94	281.0	9.02
280.0	3.31	200.0	3.18	160.0	2.27	364.1	12.8
282.3	2.93	240.0	4.21	200.4	3.83		
363.8	3.15	280.0	4.39	240.0	5.16		
		282.3	4.01	281.0	5.68		
		302.5	4.444	302.5	6.38		
		363.1	4.87	363.3	7.17		

$$f_{TGA} = 1/E_M, \tag{12}$$

where, f_{TGA} is the fitness and E_M is the least mean square error.

3. RESULTS

In this section, the proposed model is applied to the available data sets, the solubility of benzoic acid in supercritical CO_2 extraction that is shown in Table 1 (Schmitt and Reld, 1986). The extraction process performance is at a pressure range of 60-450 bar and the temperature changes among 35^oC, 45^oC, 55^oC and 70^oC. The performance of the proposed model are compared to the traditional Peng-Robinson model, lease mean square error (E_M) is selected to act as the evaluation criteria, the smaller E_M the better performance. The unit for the error E_M hereafter is $mole\ fraction^2(\times 10^{-6})$.

3.1 Periodical Changes Between Four Different Temperatures

When switching the temperature between four different values (35^oC, 45^oC, 55^oC and 70^oC) every 50 generations. The hybrid dynamic model can recognised this modification and was able to adapt to the new solution to find the suitable k_{12}.

Fig. 2 show the simulated results of the hybrid dynamic model. In Fig. 2A, The dash line represents the least mean square error of traditional model, the solid line represents the least mean square error of proposed hybrid dynamic model. Fig. 2B shows the different k_{12} values for different temperature, the dash line denotes the traditional k_{12} value used in traditional PR model, the solid line denotes the k_{12} value selected by TGA.

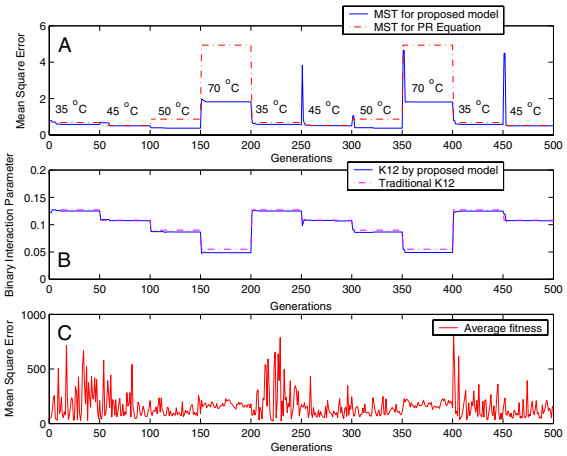

Fig. 2. Dynamic modelling for the SFE, periodical change between 35^oC, 45^oC, 55^oC and 70^oC.

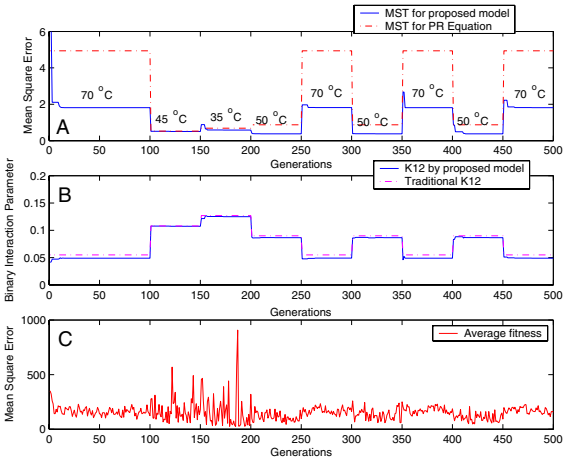

Fig. 3. Dynamic modelling for the SFE, non-periodical change between 35^oC, 45^oC, 55^oC and 70^oC.

3.2 Non-periodical Changes Between Four Different Temperatures

In this simulated study, we randomly change the temperature values of the SFE between four different values.

Fig. 3 gives out the performance of the hybrid dynamic model. In Fig. 3A, the dash line represents the least mean square error of traditional PR model, the solid line represents the least mean square error of proposed hybrid dynamic model, Fig. 3B shows the different k_{12} values for different temperature, the dash line represents the tradition k_{12} value used in traditional PR model, the solid line represents the k_{12} value selected by TGA.

As we can see that the proposed model always able to track the changes of the temperature and re-adapt to the new temperature with lower least mean square error without restart the system.

Figs. 3C and 2C give out the corresponding population analysis. Form this figure, the average fitness is far from the best fitness value, this means that the proposed system can always keep the population's diversity, this makes the system able to detect the changes and find the new solutions.

3.3 Solubility of Benzoic Acid in CO_2

Figs. 4A to 4D give the solubility of benzoic acid in supercritical CO_2 extraction, at $35°C$, $45°C$, $55°C$ and $70°C$ respectively. The experimental data are shown by stars. The results using conventional Peng-Robinson model and proposed hybrid dynamic model are shown by dashed line and solid line, respectively.

Analyse the results in Fig. 4, it shows that the proposed hybrid dynamic model can correlate the experimental data very well, the result of the proposed models is better than the traditional

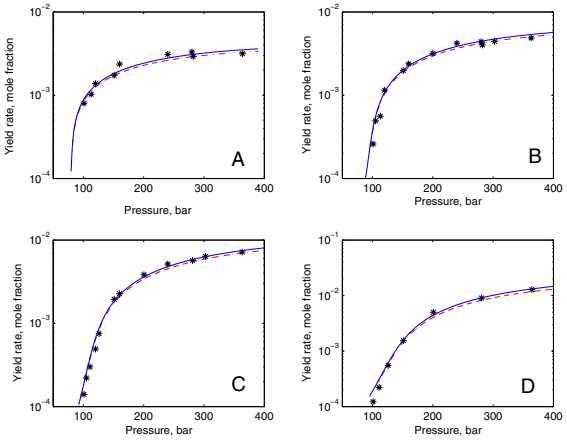

Fig. 4. Solubility of benzoic acid in CO_2 at 35^oC (A), 45^oC (B), 55^oC (C) and 70^oC (D).

Peng-Robinson equation model. The performance of the hybrid dynamic model is very well in the whole process.

In our study, the proposed hybrid dynamic model can always find the optimal value of k_{12} after the temperature changed, and get smaller mean square error than traditional models do. The proposed hybrid dynamic model for supercritical fluid extraction can successful used in the dynamic extraction process.

4. CONCLUSION

In this paper, a hybrid dynamic model combining the transformation based genetic algorithm and Peng-Robinson equation of state are proposed for dynamic modelling of the supercritical fluid extraction. The proposed hybrid dynamic model is always able to detect the change of temperature and re-adapt to the new temperature with lower least mean square error. The transformation based genetic algorithm is able to solve the

dynamic optimisation problem and produce the satisfied results. Generally the results using the proposed hybrid dynamic model is better than using the conventional Peng-Robinson model. The effectiveness of the proposed models is demonstrated by simulation and comparison studies. It is believed that the proposed approaches will be able to be applied to practical areas and highly improve the productivity.

REFERENCES

Clifford, T. (1998). *Fundamentals of Supercritical Fluids*. Oxford University Press. New York.

de Lucas, A., J. Rincon and I. Gracia (2002). Influence of operating variables on yield and quality parameters of olive husk oil extracted with supercritical carbon dioxide. *J. of the American Oil Chemists' Society* **79**(3), 237–243.

Ferreria, S. R. S, M. A. A. Meireles and F. A. Cabral (1993). Extraction of peppermint oil by supercritical carbon dioxide. *J. Food Engineering* **20**, 121–133.

Goto, M., M. Sato and T. Hirose (1993). Extraction of peppermint oil by supercritical carbon dioxide. *J. Chemical Engineering Japan* **26**, 401–407.

Kaihara, A., K. Yoshii, Y. Tsumura, S. Ishimitsu and Y. Tonogai (2002). Multiresidue analysis of 18 pesticides in fresh fruits, vegetables and rice by supercritical fluid extraction and liquid chromoatography-electrospray ionization mass spectrometry. *J. of Health Science* **48**(2), 173–178.

Li, H. and S. X. Yang (2003). Modeling of supercritical fluid extraction by hybrid peng-pobinson equation of state and genetic algorithms. *Biosystems Engineering* **86**(1), 17–25.

Li, H., S. X. Yang and J. Shi (2001). A novel neural network model for supercritical fluid extraction. In: *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*. Tucson, USA. pp. 1542–1547.

Naik, S. N., H. Lenz and R. C. Maheshvari (1989). Extraction of perfumes and flavours from plant materials with liquid carbon dioxide under liquid-vapour equilibrium conditions. *Fluid Phase Equilibria* **49**, 115–126.

Oliveira, R., M. F. Rodrigues and M. G. Bernardo (2002). Characterization and supercritical carbon dioxide extraction of walnut oil. *J. of the American Oil Chemists' Society* **79**(3), 225–230.

Schmitt, W. J. and R. C. Reld (1986). Solubility of monofunctional organic solids in chemically diverse supercritical fluids. *J. Chemical Engineering Data* **31**, 204–212.

Simöes, A. and E. Costa (2001). Using biological inspiration to deal with dynamic environments. In: *Seventh International Conference on Soft Computing*. Brno Czech Republic.

Sovava, H. (1994). Rate of the vegetable oil extraction with supercritical co_2 -i, modeling of extraction curves. *Chemical Engineering Science* **49**(3), 409–414.

Yang, S. X., H. Li and J. Shi (2003). Modeling of supercritical fluid extraction by neural networks. *Intelligent Automation and Soft Computing* **19**(1), 3–12.