HARMONY SEARCH FOR PARAMETER ESTIMATION IN VAPOR-LIQUID EQUILIBRIUM MODELING

Adrián Bonilla-Petriciolet [a], Juan Gabriel Segovia-Hernández [b] and Juan José Soto-Bernal [a]

[a] Instituto Tecnológico de Aguascalientes, Aguascalientes, México, 20256

[b] Universidad de Guanajuato, Campus Guanajuato, México, 36560

Corresponding Author's E-mail: petriciolet@hotmail.com

ABSTRACT

This study introduces the application of Harmony Search (HS) method for solving the parameter estimation problem in vapor-liquid equilibrium data modeling. The performance of this novel stochastic optimization strategy has been tested using several sets of binary vapor-liquid equilibrium data with local composition models and the classical approach of least squares. Our results indicate that HS is a promising meta-heuristic for solving global optimization problems in the modeling of phase equilibrium and may offer a better performance than those obtained using current stochastic methods such as the Simulated Annealing or the Particle Swarm Optimization.

Keywords: Global optimization, Phase equilibrium, Harmony Search, Parameter estimation.

1. INTRODUCTION

The modeling of phase equilibrium plays a major role in the design, development, operation, optimization, and control of chemical processes. A common problem for modeling phase equilibrium is to determine the parameters of a thermodynamic model used to represent a specific set of experimental information such as vapor-liquid or liquid-liquid equilibrium data. Specifically, the task is to establish the values of model parameters that provide the best fit to measured data using a proper objective function. This objective function can be formulated using either the least squares or the maximum likelihood criterion [Englezos and Kalogerakis, 2001]. For the case of vapor-liquid equilibrium (VLE) data, the parameter estimation problems usually have non-linear and non-convex solution spaces [Gau *et al*., 2000]. This optimization problem is complex in nature and difficult to solve using traditional local optimization methods due to: a) the presence of several local minima for the objective function used as optimization criterion, b) the objective function may be flat or with discontinuities in some regions of solution domain, and c) the model parameters may vary over a wide range of values. Therefore, the development of reliable methods for solving parameter estimation problems in phase equilibrium modeling is still a challenge.

Parameter estimation in VLE modeling can be performed either by direct optimization of the objective function or by solving an equivalent system of non-linear equations that is obtained from the stationary conditions of the optimization problem. Several methods for solving this thermodynamic problem are local in nature and do not guarantee convergence to the global optimum. Recently, deterministic and stochastic global optimization methods has been suggested for reliable parameter estimation in VLE modeling. Studies on deterministic optimization for phase equilibrium data modeling have been mainly focused on the

Proc. of the 5th International Symposium on Design, Operation and Control of Chemical Processes

ISBN: 978-981-08-6395-1

application of a branch and bound procedure with convex underestimating functions [Esposito and Floudas, 1998] and interval analysis using an interval-Newton/generalized bisection algorithm [Gau *et al*., 2000]. Although these methods have proven to be very promising, they are model dependent, may require problem reformulations, or significant computational time.

In the literature, several studies have shown the potential of the meta-heuristic algorithms to solve parameter estimation problems for modeling phase equilibrium data, e.g. [Alvarez *et al*., 2008; Bonilla-Petriciolet *et al*., 2010]. These methods offer several advantages for solving global optimization problems such as generality, robust performance, few information requirements for the optimization problem to be solved, easy implementation, and reasonable computational time. To date, different stochastic methods have been studied and tested for parameter estimation using VLE data, and they include: simulated annealing, genetic algorithms, differential evolution, and particle swarm optimization [Alvarez *et al*., 2008; Bonilla-Petriciolet *et al*., 2010]. These strategies usually show a robust performance but, in some challenging problems, they may fail to locate the global optimum [Bonilla-Petriciolet *et al*., 2010]. Note that the failure to find the globally optimal parameters for a thermodynamic model may cause significant errors and uncertainties in process design [Gau *et al*., 2000]. Thus, alternative optimization strategies should be studied to identify a better approach for parameter estimation problems in the context of Thermodynamics. This study introduces the application of Harmony Search (HS) for solving the parameter estimation using VLE data. HS is a new meta-heuristic that can be easily adapted for several engineering optimization problems [Lee and Geem, 2005], and our results on a variety of binary VLE data indicate that HS has a great potential for solving this and other thermodynamic calculations involved in the modeling of phase equilibrium.

2. DESCRIPTION OF HARMONY SEARCH METHOD

Harmony search is a novel meta-heuristic algorithm, which has been conceptualized using the musical process of searching for a perfect state of harmony [Geem *et al*., 2001; Lee and Geem, 2005]. This meta-heuristic is based on the analogy with music improvisation process where music players improvise the pitches of their instruments to obtain a better harmony. In the optimization context, each musician is replaced with a decision variable, and the possible notes in the musical instruments correspond to the possible values for the decision variables. The harmony in music is analogous to the optimization solution vector, and the musician's improvisations are analogous to local and global search schemes in optimization techniques. Musical performances seek to find pleasing harmony (a perfect state) as determined by an aesthetic standard, just as the optimization process seeks to find a global solution (a perfect state) as determined by an objective function [Lee and Geem, 2005].

The parameters of HS method are: the harmony memory size (HMS), the harmony memory considering rate (HMCR), the pitch adjusting rate (PAR), and the number of improvisations (NI). The harmony memory is a memory location where a set of solution vectors for decision variables is stored. The parameters HMCR and PAR are used to improve the solution vector and to increase the diversity of the search process. In HS, a new harmony (i.e., a new solution vector) is generated using three rules: 1) memory consideration, 2) pitch adjustment, and 3) random selection. It is convenient to note that the creation of a new harmony is called "improvisation". If the new solution vector (i.e., new harmony) is better than the worst one stored in HM, this new solution updates the HM. This iterative process is repeated until the given termination criterion is satisfied. Usually, the iterative steps are

performed until satisfying the following criterions: either the maximum number of succesive improvisations without improvement in the best function value, or until the maximum number of improvisations is satisfied. Figure 1 shows a flowchart of this optimization method. A full description of HS is provided by [Geem *et al*., 2001] and, in the present study, a Fortran code developed for this optimization strategy was used. This code is available to interested readers upon request to the corresponding author. It is convenient to remark that HS has been implemented in combination with a local optimization technique at the end of global search, for finding the global minimum accurately and efficiently. For local optimization, a quasi-Newton method has been used.

In this study, the suggested parameters values of HS for VLE modeling are: HMS = $10n_{var}$ where n_{var} is the number of decision variables (i.e., model parameters to be adjusted), HMCR = 0.25, and PAR = 0.75, respectively. Our preliminary calculations indicate that these parameter values are a reasonable compromise between numerical effort and reliability of HS method in this thermodynamic application. Note that the parameter NI also contributes to the trade-off between efficiency and reliability of HS. Therefore, the performance of HS has been tested by examining different levels of algorithm efficiency, which are obtained by changing the values of NI.

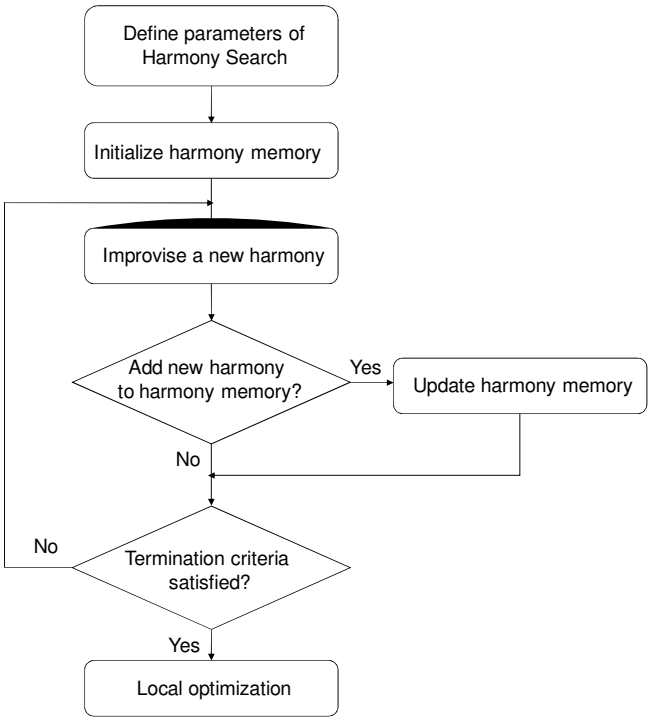

Figure 1: Flowchart of Harmony Search.

Until now, HS has been successfully applied to solve various engineering and optimization problems such as water network design, vehicle routing, soil stability analysis, heat exchanger design, and transportation energy modeling [Lee and Geem, 2005]. In the context of chemical engineering, there are few studies concerning the application of this novel

stochastic method and, to the best of our knowledge, the performance of HS in parameter estimation for phase equilibrium modeling has not yet been reported. Thus, this study introduces the application of HS for solving this challenging thermodynamic problem. In the following section, the formulation of the global optimization problem is described.

3. THE PARAMETER ESTIMATION PROBLEM FOR VAPOR-LIQUID EQUILIBRIUM MODELING

As indicated, the objective function used for parameter estimation problem in VLE data modeling can be derived from either the least squares principle or the maximum likelihood criterion [Gau *et al*., 2000; Englezos and Kalogerakis, 2001]. In the first approach, the model parameters are determined whereas, in the second formulation, both the true values of state variables and model parameters are obtained. In this study, we have used the least squares approach for illustrating the performance of HS as global optimization strategy in the modeling of VLE data.

In the least squares formulation, the objective function is defined as the difference between the experimental data and the calculated values using a proper model. Literature indicates that the least squares principle is the most popular approach for obtaining parameters of thermodynamic models for VLE and other thermodynamic data [López *et al*., 2006]. For the case of VLE data (i.e., *x*–*y*–*P* at constant *T*, or *x*–*y*–*T* at constant *P*), excess Gibbs energy equations are widely employed for phase equilibrium modeling. The objective function commonly used for data fitting of VLE data is based on activity coefficients and is usually defined as

$$F_{obj} = \sum_{j=1}^{ndat}\sum_{i=1}^{c}\left(\frac{\gamma_{ij}^{\exp} - \gamma_{ij}^{calc}}{\gamma_{ij}^{\exp}}\right)^2 \tag{1}$$

where $\gamma_i^{\exp}$ and γ_i^{calc} are respectively the experimental and calculated values for the activity coefficient of component *i*, *ndat* is the number of VLE experimental data, and *c* is the number of components in the mixture, respectively.

The equilibrium between vapor and liquid phases in a multi-component system implies that *T*, *P* and the fugacities of each component must be the same in both phases. At low pressure, the fugacity coefficient of pure components nearly cancels each other and Poynting corrections usually are very close to unity. With these assumptions, $\gamma_i^{\exp}$ can be calculated from VLE data using the following expression

$$\gamma_i^{\exp} = \frac{y_i^{\exp} P}{x_i^{\exp} P_i^0} \quad i = 1,...,c \tag{2}$$

where $x_i^{\exp}$ and $y_i^{\exp}$ are respectively the experimental mole fractions of component *i* in liquid and vapor phases at equilibrium, and P_i^0 is the vapor pressure of pure component *i* at the system temperature *T*. In our study, the Wilson, NRTL and UNIQUAC models are used to calculate the activity coefficients, γ_i^{calc}, and in parameter estimation we seek the numerical values of the energy parameters of these models that will minimize Eq. (1). The energy parameters of these thermodynamic models are defined as in the DECHEMA. It is important to note that the global minimization of least squares objective function can be done as an

unconstrained optimization problem using these local composition models. However, several studies have shown that, even for relatively simple thermodynamic equations such as Wilson model, multiple local optima can occur in parameter estimation for VLE data modeling [Gau *et al*., 2000; Alvarez *et al*., 2008; Bonilla-Petriciolet *et al*., 2010]. This is because the highly non-linear form of the thermodynamic models makes F_{obj} strongly non-linear, potentially non-convex with several local minima within the specified bounds.

The performance of HS has been tested using several sets of binary vapor-liquid equilibrium data. Table 1 provides the details (i.e., conditions of experimental data, thermodynamic models, and global optimum) of all VLE examples used to evaluate the HS method. All the experimental data are taken from DECHEMA collection and the vapor pressure is calculated by Antoine equation using the parameters reported in [Gau *et al.*, 2000]. These VLE problems have been studied using both deterministic and stochastic optimization methods, e.g. [Esposito and Floudas, 1998; Gau *et al*., 2000; Alvarez *et al*., 2008; Bonilla-Petriciolet *et al*., 2010]. According to reported studies, there are at least two local minima (including the global minimum) in the specified interval for decision variables of all VLE problems. Therefore, we consider that the number and features of these problems are sufficient to demonstrate and compare the performance of HS method.

Table 1: VLE problems selected to test the performance of Harmony Search method

System	*P* or *T*	Model	Global optimum
tert butanol + 1 butanol	100 mm Hg	Wilson	0.01026
	500 mm Hg		0.00685
	700 mm Hg		0.01369
water + 1,2 ethanediol	430 mm Hg	Wilson	1.03913
		UNIQUAC	1.40855
		NRTL	1.25353
benzene + hexafluorobenzene	30 °C	Wilson	0.01178
	50 °C		0.00894
	300 mm Hg		0.01486
	760 mm Hg		0.01462

4. RESULTS AND DISCUSSION

All VLE examples are solved 100 times each, starting from a different, random point inside specified bounds on decision variables, and the performance of HS method is tested based on both reliability (measured in terms of number of times the algorithm located the global minimum out of 100 trials, refereed as success rate SR) and computational efficiency (measured in terms of average number of function evaluations NFE). The average NFE is evaluated using successful trials only, where a trial is considered successful if the global optimum is obtained with an absolute error of 10^{-5} or less in the objective function value. For the sake of brevity, algorithm reliability results are summarized through the global success rate (*GSR*). Specifically, *GSR* is defined as the percent of successes out of all calculations performed on the collection of VLE problems tested

$$GSR = \frac{1}{N_{prob}} \sum_{i=1}^{N_{prob}} SR_i \tag{3}$$

where SR_i is the success rate in each problem tested. We optimize Eq. (1) with respect to the parameters of thermodynamic models inside the interval: $\theta_1, \theta_2 \in$ (-10,000, 20,0000) for Wilson model, $\theta_1, \theta_2 \in$ (-5,000, 20,000) for UNIQUAC model, and $\theta_1, \theta_2 \in$ (-2,000, 5,000) and $\alpha_{12} \in$ (0.01, 10.0) for NRTL model where θ_1 and θ_2 are the corresponding energy parameters. As stated, the initial values for each calculation are randomly generated within these bounds.

With illustrative purposes, the results of solving the VLE parameter estimation problems for different values of NI (i.e., NFE) are shown in Figure 1. To directly test and compare the performance of HS with those obtained for other stochastics methods, we keep their numerical effort the same via NFE and analyze the results obtained in terms of GSR. In our analysis, the results from [Bonilla-Petriciolet *et al*., 2010] reported for Simulated Annealing (SA), Particle Swarm Optimization (PSO) and Differential Evolution (DE) are used to compare with those of HS method. Therefore, the performance of these methods are also included in Figure 1.

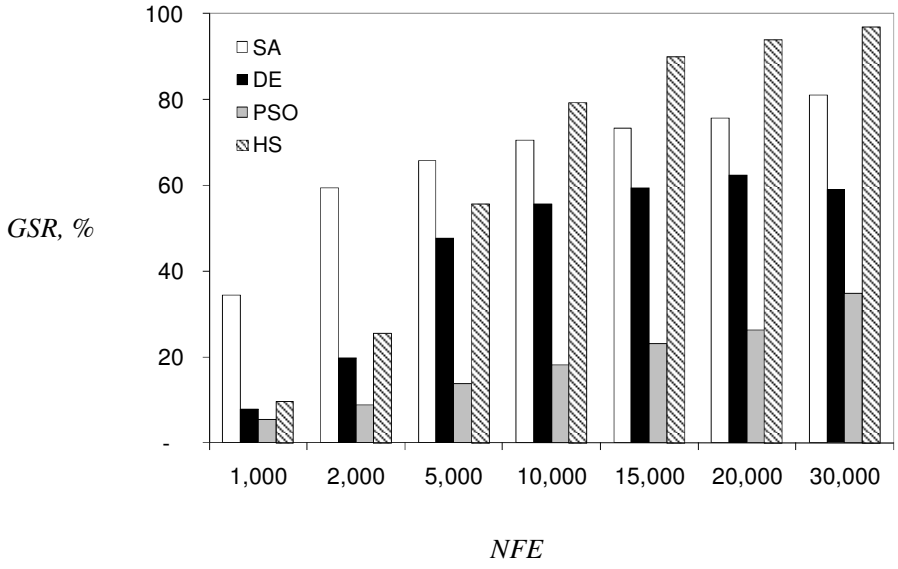

Figure 2: Global success rate (*GSR*) *versus* average number of function evaluations (*NFE*) of Simulated Annealing (SA), Differential Evolution (DE), Particle Swarm Optimization (PSO) and Harmony Search (HS) for VLE data modeling.

Our results indicate that HS is a promising direct-search method for solving global optimization problems in the modeling of phase equilibrium and may offer a better performance than those reported for other stochastic methods if proper numerical effort is allowed (see Figure 1). As expected, the reliability of the stochastic methods highly depends on and increases with NFE. In general, all the stochastic methods may fail in the global minimization of the objective functions involved in parameter estimation for VLE data modeling. It appears that these methods may be trapped by the local minima of the objective

functions of selected VLE problems especially at early NFE. This could be because all minima found for VLE parameter estimation problems generally are located in a relatively narrow valley in the parameter space and also these minima may be comparable in some problems. In particular, the VLE problems tested involve challenging global optimization functions and illustrate the difficulty of locating the globally optimal parameters for VLE data modeling even using simple thermodynamic models.

The reliability of HS is better than that of DE, PSO and SA for tested VLE examples using NFE > 10,000 and the GSR of HS ranged from 10 to 97%. If a low numerical effort is considered, SA outperformed other stochastic methods for solving VLE problems in terms of global success rate. The GSR of SA ranged from 34 to 81% in VLE problems. On the other hand, DE can achieve a maximum GSR of 59%, while the maximum GSR obtained for PSO was 35% for this set of VLE problems. Specifically, our results indicate that the GSR of SA is better than those obtained for DE, PSO and HS in early iterations. However, as NFE increases, the performance of HS improves and its reliability is higher than those reported for SA. These results suggest that SA may be preferred for parameter estimation in the VLE modeling if a small number of function evaluations is allowed; otherwise, HS is more reliable.

In summary, these results indicate that the GSR of HS and SA is better compared to DE and PSO for tested values of NFE. The computing time of stochastic methods increases directly proportional to problem dimensionality (i.e., number of decision variables). However, compared to other stochastic methods, HS imposes fewer mathematical requirements and its implementation is straightforward. So, HS is among the best algorithms and can be considered an alternative method for parameter estimation in VLE modeling.

5. CONCLUSIONS

In this study, the performance of HS has been tested and compared for VLE modeling using experimental data for binary systems and least squares criterion. Overall, HS was found to perform better than other algorithms in terms of success rate for parameter estimation in VLE data modeling if a proper numerical effort is used. The structure of HS algorithm is relatively easier and favors its hybridization with other stochastic methods. In fact, the hybridization of HS using other meta-heuristic is an atractive alternative for improving the performance of HS especially for challenging optimization problems involved in the modeling of phase equilibrium.

ACKNOWLEDGEMENTS

We acknowledge the financial support provided by CONACyT, DGEST, Instituto Tecnológico de Aguascalientes and Universidad de Guanajuato.

REFERENCES

1. Alvarez V.H., Larico R., Ianos Y. and Aznar M. Parameter estimation for VLE calculation by global minimization: the genetic algorithm. *Brazilian Journal of Chemical Engineering,* **25**, 409-418, 2008.
2. Bonilla-Petriciolet A., Bravo-Sánchez U.I., Castillo-Borja F., Zapiain-Salinas J.G. and Soto-Bernal J.J. The performance of simulated annealing in parameter estimation for vapor-liquid equilibrium modeling. *Brazilian Journal of Chemical Engineering*, **24**, 151-162, 2007.

3. Bonilla-Petriciolet A., Rangaiah G.P. and Segovia-Hernández J.G. Evaluation of stochastic global optimization methods for modeling vapor-liquid equilibrium data. *Fluid Phase Equilibria*, **287**, 111-125, 2010.
4. Englezos P. and Kalogerakis N. *Applied parameter estimation for chemical engineers*, Marcel-Dekker, New York, 2001.
5. Esposito W.R. and Floudas C.A. Global optimization in parameter estimation of nonlinear algebraic models via the error-in-variables approach. *Industrial Engineering Chemistry Research*, **37**, 1841-1858, 1998.
6. Gau C.Y., Brennecke J.F. and Stadtherr M.A. Reliable parameter estimation in VLE modeling. *Fluid Phase Equilibria*, **168**, 1-18, 2000.
7. Geem Z.W., Kim J.H. and Loganthan G.V. A new heuristic optimization algorithm: harmony search. *Simulation*, **76**, 60-68, 2001.
8. Lee K.S. and Geem Z.W. A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice. *Computer Methods in Applied Mechanics and Engineering*, **194**, 3902-3933, 2005.
9. López J.A., Trejos V.M. and Cardona C.A. Objective functions analysis in the minimization of binary VLE data for asymmetric mixtures at high pressures. *Fluid Phase Equilibria*, **248**, 147-157, 2006.

Brief Biography of the Presenter

Adrián Bonilla-Petriciolet is Professor in the Department of Chemical Engineering of Instituto Tecnológico de Aguascalientes, México. His research interests are applied thermodynamics, process design and optimization. He has published several papers in international journals and presented a number of papers in conferences.