

THE PERFORMANCE OF SIMULATED ANNEALING IN PARAMETER ESTIMATION FOR VAPOR-LIQUID EQUILIBRIUM MODELING

A. Bonilla-Petriciolet^{1*}, U. I. Bravo-Sánchez¹, F. Castillo-Borja¹,
J. G. Zapiain-Salinas¹ and J. J. Soto-Bernal²

¹Departamento de Ingeniería Química, Instituto Tecnológico de Aguascalientes,
Phone and Fax: (52) 449 910-5002 ext. 127, CP: 20256, Aguascalientes, Aguascalientes, México.
Email: petriciolet@hotmail.com

²Departamento de Sistemas y Computación Instituto Tecnológico de Aguascalientes,
CP: 20256, Aguascalientes, Aguascalientes, México.

(Received: May 30, 2005 ; Accepted: December 4, 2006)

Abstract - In this paper we report the application and evaluation of the simulated annealing (SA) optimization method in parameter estimation for vapor-liquid equilibrium (VLE) modeling. We tested this optimization method using the classical least squares and error-in-variable approaches. The reliability and efficiency of the data-fitting procedure are also considered using different values for algorithm parameters of the SA method. Our results indicate that this method, when properly implemented, is a robust procedure for nonlinear parameter estimation in thermodynamic models. However, in difficult problems it still can converge to local optimums of the objective function.

Keywords: Vapor-liquid equilibrium; Simulated annealing; Nonlinear parameter estimation; Error-in-variable method; Global optimization.

INTRODUCTION

In many important areas of chemical, biochemical and petroleum engineering, mathematical models form the basis for the design, optimization and control of process systems (Esposito and Floudas, 1998; Englezos and Kalogerakis, 2001). Parameter estimation is a common problem in many engineering applications and is one of the steps involved in the formulation and validation of a mathematical model that describes a process of interest. Specifically, it refers to the process of obtaining values of the parameters by matching the model-based calculated values with the set of measurements (Englezos and Kalogerakis, 2001). If we use linear model equations, the problem is called linear estimation, while nonlinear estimation refers

to the more general problem where the model equations are nonlinear functions of the parameters. As indicated by Esposito and Floudas (1998), the use of nonlinear models introduces an added level of complexity into the numerical estimation of model parameters.

Parameter estimation is generally based on the classical least squares or maximum likelihood approaches (Esposito and Floudas, 1998; Gau et al., 2000; Gau and Stadtherr, 2002). In the classical least squares approach, it is assumed that there is a set of independent variables not subject to measurement error, while errors in all measured variables are accounted for in the maximum likelihood approach. Both methods involve the minimization of an objective function subject to constraints representing the model equations. Depending on the system

*To whom correspondence should be addressed

characteristics, we can solve a constrained or unconstrained optimization problem, which in general is nonlinear and potentially nonconvex.

Several local and global optimization methods have been used for parameter estimation in both linear and nonlinear models (Bard, 1974; Britt and Luecke, 1973; Fabries and Renon, 1975; Gmehling et al., 1977-1990; Anderson et al., 1978; Valko and Vajda, 1987; Tjoa and Biegler, 1992; Vamos and Hass, 1994; Esposito and Floudas, 1998; Gau et al., 2000; Gau and Stadtherr, 2002). Local optimization methods are not reliable for finding the global minimum of the problem; however, these methods are the most widely used. In the other hand, some applications of global optimization methods in parameter estimation have been reported by Esposito and Floudas (1998), Kleiber and Axmann (1998), Park and Froment (1998), Gau et al. (2000), Costa et al. (2000), Gau and Stadtherr (2002) and Dominguez et al. (2002).

An attractive alternative for the reliable solution of nonlinear parameter estimation is the use of stochastic optimization methods such as simulated annealing (SA) or genetic algorithm (GA). Even though these methods provide no formal guarantee for global optimization, they are reliable strategies and offer a reasonable computational effort in the optimization of multivariable functions (Michalewicz and Fogel, 1999). In fact, some papers have reported the application of SA and GA in parameter estimation using chemical engineering models (Kleiber and Axmann, 1998; Park and Froment, 1998; Costa et al., 2000).

In the context of thermodynamics, parameter estimation may also encounter computational difficulties due to the possibility of several local optimums in the objective function used as optimization criterion. As indicated by Dominguez et al. (2002), failing to identify the global optimum in parameter estimation may cause errors and uncertainties in equipment design and erroneous conclusions about model performance. In view of this, it is important to test different optimization methods to identify a reliable and efficient technique for this purpose.

In this paper, we report the application and evaluation of the SA method in parameter estimation for vapor-liquid equilibrium (VLE) modeling. We have tested the numerical performance of SA using the classical least squares and maximum likelihood approaches. Reliability and efficiency in the data-fitting procedure are also considered using different values for the algorithm parameters of the SA method.

METHODOLOGY

Problem Formulation

In accordance with Gau et al. (2000), consider that a set of observations y_{ij} of $i=1, \dots, m$ dependent response variables from $j=1, \dots, \text{ndat}$ experiments is available, where the responses will be adjusted to an explicit model $y_{ij} = f_i \left(\vec{x}_j, \vec{\theta} \right)$ with independent variables $\vec{x}_j = (x_{1j}, \dots, x_{pj})^T$ and parameters $\vec{\theta} = (\theta_1, \dots, \theta_q)^T$. Measurement errors in \vec{x}_j can either be treated or neglected, and depending on the choice, we can have a least squares or maximum likelihood formulation. Different objective functions can be used to obtain the parameter values that provide the best fit. For the case of the classical least squares (LS) criterion, we use the following objective function

$$F_{\text{obj}} = \sum_{j=1}^{\text{ndat}} \sum_{i=1}^m \left(\frac{y_{ij} - f_i \left(\vec{x}_j, \vec{\theta} \right)}{y_{ij}} \right)^2 \quad (1)$$

This function is optimized with respect to the model parameters $\vec{\theta} = (\theta_1, \dots, \theta_q)^T$ and Gmehling et al. (1977-1990), Gau et al. (2000) and Dominguez et al. (2002) have used this formulation in data fitting for VLE modeling.

If we assume that there are measurement errors in the state variables z_{ij} for the experiments in the system to be modeled, the error-in-variable (EIV) formulation has the form

$$F_{\text{obj}} = \sum_{j=1}^{\text{ndat}} \sum_{i=1}^{\text{nest}} \frac{(z_{ij}^t - z_{ij})^2}{\sigma_i^2} \quad (2)$$

subject to

$$\vec{g} \left(z_{ij}^t, \vec{\theta} \right) = 0 \quad i = 1, \dots, \text{nest} \quad j = 1, \dots, \text{ndat} \quad (3)$$

where \vec{g} is a vector of np model functions, nest is

the number of state variables, z_{ij}^t represents the unknown “true” values of state variables for each measurement and σ_i represents the standard deviation associated with the measurement of state variable i . The optimization variables of this problem are the set of z_{ij}^t and the model parameters $\vec{\theta} = (\theta_1, \dots, \theta_q)^T$. In this case, there is a substantial increase in the dimensionality of the optimization problem, which depends on the number of experiments. Esposito and Floudas (1998) and Gau and Stadtherr (2002) have applied their deterministic optimization approaches to solving Eqs. [2] and [3] with different chemical engineering models including vapor-liquid equilibrium equations.

The last functions can be minimized either as a constrained or as an unconstrained optimization problem. In all our examples, we will consider only the unconstrained formulation. We have studied the numerical performance of the SA method in the global minimization of these functions using vapor-liquid equilibrium models. In the next section, we describe the algorithm used for the SA method.

Simulated Annealing (SA)

Simulated annealing is a stochastic optimization technique inspired by the thermodynamic process of cooling of molten metals to attain the lowest free energy state (Kirkpatrick et al., 1983). Starting with an initial solution and armed with adequate perturbation and evaluation functions, the algorithm performs a stochastic partial search of the state space. In minimization problems, uphill moves are occasionally accepted with a probability controlled by a parameter called annealing temperature, T_{SA} . The probability of acceptance of uphill moves decreases as T_{SA} decreases. At high temperatures, the search is almost random, while at low temperatures the search becomes almost greedy. At zero temperature, the search becomes totally greedy, i.e., only good moves are accepted (Kirkpatrick et al., 1983). The core of the algorithm is the Metropolis procedure, which simulates the annealing process at a given T_{SA} (Metropolis et al., 1953). The Metropolis criterion is used to accept or reject the

uphill moves. Several algorithms have been proposed for the SA method. We have used the algorithm proposed by Corana et al. (1987) because previous papers have reported its reliability and efficiency in thermodynamic calculations (Henderson et al., 2001; Rangaiah, 2001).

In the algorithm proposed by Corana et al. (1987), a trial point is randomly chosen within the step length VM (a vector of length n variables) of a starting point. The function is evaluated at this trial point and its value is compared to its value at the initial point, where the Metropolis criterion is used to accept or reject the trial point. If the trial point is accepted, the algorithm moves on from that point. If it is rejected, another point is chosen instead for a trial evaluation. Each element of VM is periodically adjusted so that half of all function evaluations in that direction are accepted. A fall in T_{SA} , after NT iterations, is imposed upon the system with the RT variable by

$$T_{SAj+1} = RT \times T_{SAj} \quad (4)$$

where j is the iteration counter and RT is the temperature reduction factor. For our examples, we have defined this parameter equal to 0.85 which is the value suggested by Goffe et al. (1994). Thus, as T_{SA} declines, downhill moves are less likely to be accepted and the percentage of rejections rises. Given the scheme for the selection of VM, VM falls. Thus, as T_{SA} declines, VM falls and SA focuses upon the most promising area for optimization. A full description of this algorithm is found in Corana et al. (1987) and we have used the subroutine developed by Goffe et al. (1994). Figure 1 shows a flow diagram of this optimization method. The choice of cooling schedule is a crucial aspect in the implementation of SA because it affects the numerical performance of the optimization procedure. Considering this fact, we have tested the effect of parameters T_{SA} and NT on the numerical performance of SA to find the global optimum of the objective function used in the data-fitting procedure. Several calculations are done to perform this evaluation and a statistical analysis is used to establish their relative effect on the reliability of the SA method.

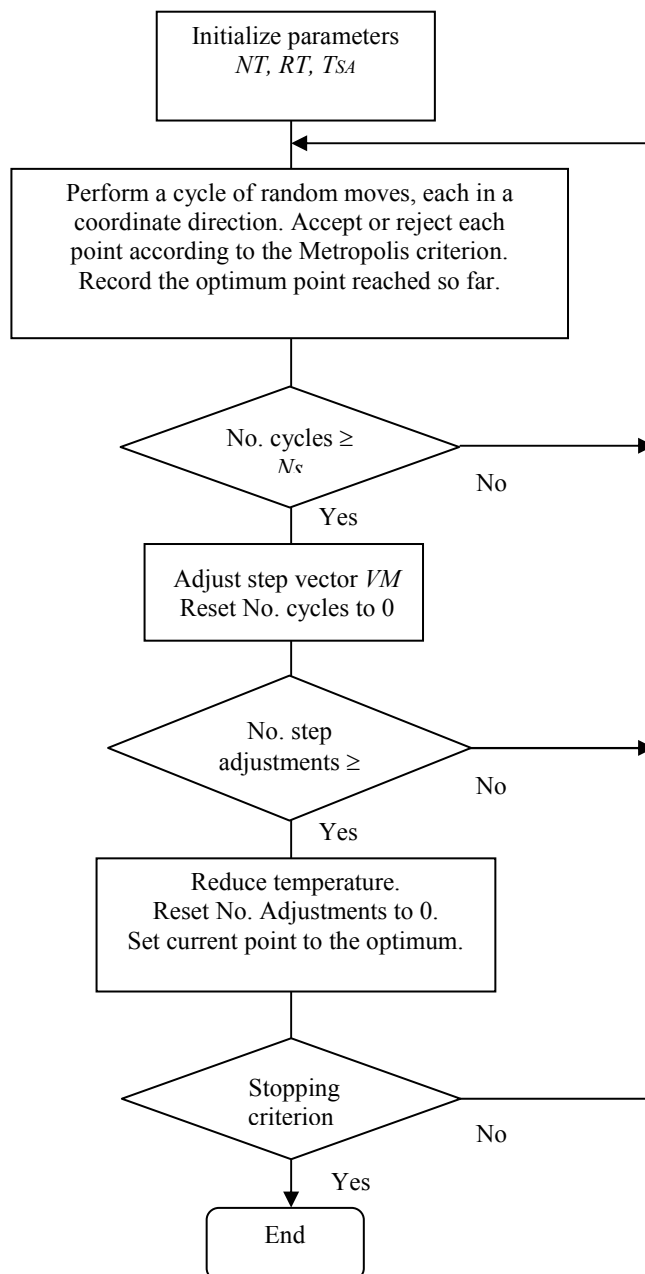


Figure 1: Flow diagram of SA optimization algorithm (Corana et al., 1987)

RESULTS AND DISCUSSION

We test the SA method in the data-fitting procedure using several experimental VLE data and different thermodynamic models. All systems used in this paper have an objective function with multiple optimums and they have been applied for testing global deterministic optimization methods (Esposito and Floudas, 1998; Gau et al., 2000; Gau and Stadtherr, 2002). The reliability and efficiency of SA is evaluated based on the following criteria: a) success rate in finding the global minimum SR and b) average number of objective function evaluations

during the optimization procedure NFEV. All examples are solved 100 times (each time with a different random initial value and random number seed) and the results obtained are divided into four groups (each group with 25 calculations). For each group of results, SR and NFEV are calculated and a variance analysis is performed to establish the effect of parameters T_{SA} and NT on the reliability of the data-fitting procedure. We consider the statistical effect of these parameters significant if the value of the p-level obtained from the statistical analysis is lower than 0.05. In all our examples, we have used a tolerance of $1.0E-09$ for the convergence of the SA

algorithm. The experimental data are taken from DECHEMA and Reid et al. (1987). In all examples, we have assumed gas ideal behavior and the vapor pressures of pure components are calculated using the Antoine equation. For problems 1 – 4, we use the constants reported in DECHEMA for the Antoine equation.

Problem 1. Tert Butanol-1 Butanol

Our first example is the VLE for the binary system tert butanol-1 butanol. This system was studied by Gau et al. (2000) using an interval analysis approach and classical least square formulation. We use three sets of experimental data under different isobaric conditions. The Wilson equation is used to calculate the liquid-phase activity coefficients, which are defined by

$$\ln \gamma_1 = -\ln(x_1 + \Lambda_{12}x_2) + x_2 \left[\frac{\Lambda_{12}}{x_1 + \Lambda_{12}x_2} - \frac{\Lambda_{21}}{\Lambda_{21}x_1 + x_2} \right] \quad (5)$$

$$\ln \gamma_2 = -\ln(x_2 + \Lambda_{21}x_1) - x_1 \left[\frac{\Lambda_{12}}{x_1 + \Lambda_{12}x_2} - \frac{\Lambda_{21}}{\Lambda_{21}x_1 + x_2} \right] \quad (6)$$

The binary parameters Λ_{12} and Λ_{21} are given by

$$\Lambda_{12} = \frac{v_2}{v_1} \exp\left(-\frac{\theta_1}{RT}\right) \quad (7)$$

$$\Lambda_{21} = \frac{v_1}{v_2} \exp\left(-\frac{\theta_2}{RT}\right) \quad (8)$$

where v_1 and v_2 are the pure component liquid molar volumes, T is the system temperature and θ_1 and θ_2 are the energy parameters. Assuming an ideal vapor phase, the experimental values for the activity coefficients γ_i^{exp} are calculated using the next equation

$$\gamma_i^{\text{exp}} = \frac{y_i^{\text{exp}} P^{\text{exp}}}{x_i^{\text{exp}} P_i^0} \quad i = 1, \dots, c \quad (9)$$

where P_i^0 is the vapor pressure of pure component i at the system temperature T and c is the number of components. The vapor pressure is calculated using

$$\log_{10} P_i^0 = a_1 - \frac{b_1}{c_1 + T} \quad (10)$$

where T is in °C and P_i^0 , in mmHg. In accordance with Gmehling (1977-1990) and Gau et al. (2000), we use a relative least squares formulation to fit the data

$$F_{\text{obj}} = \sum_{j=1}^{\text{ndat}} \sum_{i=1}^c \left(\frac{\gamma_{ij}^{\text{exp}} - \gamma_{ij}^{\text{calc}}}{\gamma_{ij}^{\text{exp}}} \right)^2 \quad (11)$$

We optimize this function with respect to the Wilson model parameters inside the intervals $\theta_1 \in (-8500, 320000)$ and $\theta_2 \in (-8500, 320000)$. The initial values for each calculation are randomly generated within these intervals. The performance of SA is tested using the next arbitrary values for its parameters: $T_{\text{SA}}(10, 100, 1000)$ and $\text{NT}(10, 20)$. The results obtained by other thermodynamic calculations suggest that our proposed values for T_{SA} and NT are suitable choices and favor the numerical performance of SA. We don't use higher values of NT because the computational time of the SA method is significantly increased. The results obtained for all combinations of $T_{\text{SA}} - \text{NT}$ are reported in Table 1 and the statistical effect of these parameters on the reliability of the data-fitting procedure is reported in Table 2. For all pressure conditions, SA fails several times to find the global minimum of the objective function. In all data sets, the parameter T_{SA} shows a significant statistical effect on the reliability of the parameter estimation procedure ($p\text{-level} \leq 0.05$). In general, an increase in T_{SA} improves the reliability of the SA method, of course at the expense of a major computational effort. Only for the experimental data at 700 mmHg, the parameter NT and the interaction $T_{\text{SA}} - \text{NT}$ show a significant statistical effect on the reliability of the data fitting. The global optimums reported in Table 1 are consistent with the results of Gau et al. (2000).

Table 1: Results of parameter estimation for the tert butanol-1 butanol system using simulated annealing with the least squares formulation

P, mm Hg	Number of data	Global minimum ²			NT	NFEV(SR,%) ¹		
		θ_1	θ_2	F_{obj}		T_{SA}		
						10	100	1000
100	9	-568	745.3	0.0103	10	57005±1389.7 (67.0±6.8)	62485±1375.4 (87.0±10.0)	68289±1555.0 (90.0±4.0)
					20	114953±2853.3 (80.0±17.0)	126505±2634.1 (92.0±3.3)	137161±2910.9 (88.0±5.7)
500	9	-718	1264.7	0.0069	10	57709±1522.2 (74.0±6.9)	63325±1377.1 (92.0±6.5)	68853±1458.0 (90.0±5.2)
					20	115529±2542.8 (84.0±7.3)	127065±2813.1 (91.0±3.8)	138305±2553.4 (87.0±6.8)
700	9	-734	1318.2	0.0137	10	57773±1321.5 (70.0±6.9)	63349±1374.7 (97.0±3.8)	68937±1466.4 (94.0±7.7)
					20	115393±2632.8 (95.0±7.6)	127049±2777.6 (95.0±5.0)	137793±3089.2 (93.0±6.8)

¹NFEV is defined as the average number of objective function evaluations during the optimization procedure and SR is the success rate in finding the global minimum based on 4 groups of 25 calculations with random initial values and number seeds.

²Global optimums are consistent with results reported by Gau et al. (2000).

Table 2: Results of variance analysis for the data-fitting procedure using simulated annealing in vapor-liquid modeling

System	Formulation	ELV Model	T, °C or P, mm Hg	Variance analysis, p-level ¹		
				T_{SA}	NT	Interaction T_{SA} - NT
Tert butanol-1 Butanol	LS	Wilson-Ideal gas	P=100	0.0032	0.1668	0.2793
			P=500	0.0021	0.4410	0.1072
			P=700	0.0013	0.0124	0.0007
Water-1,2 Ethanediol	LS	Wilson-Ideal gas	P=430	0.0000	0.0547	0.0314
		NRTL-Ideal gas				
		UNIQUAC-Ideal gas		0.0000	0.0010	0.0095
Benzene-Hexafluorobenzene	LS	Wilson-Ideal gas	T=30	0.0620	0.0231	0.1568
			T=50	0.0020	0.0077	0.0020
			P=300	0.0705	0.0064	0.4290
			P=760			
Benzene-Hexafluorobenzene	EIV	Wilson-Ideal gas	T=50	0.0380	0.0781	0.1016
			T=70	0.0288	0.0954	0.3007
			P=300	0.0290	0.2727	0.4398
			P=760	0.0414	0.0000	0.1709
Methanol-1,2 Dichloroethane	EIV	van Laar-Ideal gas	T=50			

¹Statistical effect of SA parameters on the reliability of the data-fitting procedure. The variable SR is used as dependent variable in the variance analysis.

Problem 2. Water-1,2 Ethanediol

We use one data set for the system water-1,2 ethanediol at 430 mmHg. This system was also analyzed by Gau et al. (2000) using the least squares formulation and they reported several local optimums for the objective function. Also, they indicated that global optimal parameters provide a better fit to the experimental data. We optimize Eq. 11 with respect to the energy parameters of the Wilson, NRTL and UNIQUAC models. The energy parameters for the NRTL and UNIQUAC equations are defined as in DECHEMA. We use the same initial intervals employed in the last example for the Wilson model, while the feasible domain for the NRTL and UNIQUAC parameters are $\theta_1, \theta_2 \in (-2000, 5000)$ and

$\alpha_{12} \in (0.01, 10)$ and $\theta_1, \theta_2 \in (-5000, 20000)$, respectively. These domains include the parameters reported in DECHEMA. Our results are presented in Table 3 and the statistical analysis is reported in Table 2. For the Wilson and UNIQUAC models, the parameter T_{SA} and the interaction $T_{SA} - NT$ show a significant statistical effect on the reliability of the data-fitting procedure. Also, SA shows some failures in the global minimization of the objective function. The global minimum reported for the Wilson model is in agreement with that published by Gau et al. (2000). The energy parameters reported in DECHEMA correspond to local optimums for the three models. Similar findings with other systems have been reported by Dominguez et al. (2002) using the Wilson and UNIQUAC equations.

Table 3: Results of parameter estimation for the water-1,2 ethanediol system at 430 mmHg using simulated annealing with the least squares formulation

Model	Number of data	Global minimum ²				NT	NFEV(SR,%) ¹		
		θ_1	θ_2	α_{12}	F_{obj}		T_{SA}		
							10	100	1000
Wilson-Ideal gas	18	5072.4	-1921.6		1.0391	10	57697±1620.2 (64.0±13.5)	63237±1758.9 (99.0±2.0)	69053±1335.3 (100.0±0.0)
						20	115737±3068.1 (80.0±7.3)	127569±2437.2 (99.0±2.0)	138537±3132.7 (100.0±0.0)
NRTL-Ideal gas		-679.0	3046.1	0.6214	1.2535	10	88177±1866 (100±0.0)	96883±1757 (100±0.0)	105529±1562 (100±0.0)
						20	176509±3467 (100±0.0)	194377±2839 (100±0.0)	210301±3485 (100±0.0)
UNIQUAC-Ideal gas		-1131.8	3617.7		1.4085	10	57685±1622 (58.0±10.0)	63537±1513 (94.0±5.0)	69357±1285 (99.0±2.0)
						20	115113±2778 (78.0±8.0)	126969±2604 (100.0±0.0)	138049±2737 (100.0±0.0)

¹NFEV is defined as the average number of objective function evaluations during the optimization procedure and SR is the success rate in finding the global minimum based on 4 groups of 25 calculations with random initial values and number seeds.

²Global optimum is consistent with the results reported by Gau et al. (2000).

Problem 3. Benzene-Hexafluorobenzene (LS Formulation)

The parameter estimation of this problem is performed assuming VLE data under isothermal and isobaric conditions. This system has also been analyzed by Gau et al. (2000) using interval analysis. Four data sets are used to test the numerical performance of the SA method. We use a least squares formulation in the parameter estimation and the same initial intervals as those for the energy parameters of the Wilson model are applied. The results of parameter estimation are reported in Table 4, while the statistical analysis appears in Table 2. In this example, SA has a better numerical performance; however, it still shows some failures in the global minimum of the objective function. Only

for the data at 760 mmHg, SA finds the global minimum with a reliability of 100% and its numerical performance is independent of the values for parameters T_{SA} and NT. We find a significant statistical effect of parameter NT in the reliability of the parameter estimation procedure for the data at 30 °C, 50 °C and 300 mmHg. For the case of T_{SA} , it affects the reliability of SA for data at 50 °C. As indicated by Gau et al. (2000), a failure in the global minimization in parameter estimation affects the capability of the thermodynamic model in prediction of homogeneous azeotropes in this system. That failure in the model's performance in the calculation of azeotropes could result in design problems in separation processes. The global optimums found with SA are in agreement with those reported by Gau et al. (2000).

Table 4: Results of parameter estimation for the benzene-hexafluorobenzene system using simulated annealing with the least squares formulation

T, °C or P, mm Hg	Number of data	Global minimum ²			NT	NFEV(SR, %) ¹		
		θ_1	θ_2	F_{obj}		T_{SA}		
						10	100	1000
T = 30	10	-467.8	1313.9	0.0118	10	57841±1345.4 (91.0±8.9)	63569±1523.1 (96.0±3.3)	69081±1308.2 (100.0±0.0)
					20	116105±2559.4 (99.0±2.0)	126833±2581.5 (100.0±0.0)	138449±3124.0 (100.0±0.0)
50	11	-424.1	983.1	0.0089	10	57745±1216.1 (97.0±2.0)	63413±1284.5 (100.0±0.0)	68957±1588.7 (100.0±0.0)
					20	115929±2647.4 (100.0±0.0)	127337±2446.8 (100.0±0.0)	138649±2329.5 (100.0±0.0)
P = 300	17	-432.5	992.9	0.0149	10	58097±1315.2 (89.0±6.0)	63501±1264.1 (95.0±5.0)	69177±1277.4 (96.0±3.3)
					20	115505±2873.8 (97.0±3.8)	126937±2881.2 (99.0±2.0)	138481±2868.7 (99.0±2.0)
760	29	-334.7	704.7	0.0146	10	57901±1262.8 (100.0±0.0)	63641±1270.6 (100.0±0.0)	69217±1378.5 (100.0±0.0)
					20	115673±2580.7 (100.0±0.0)	126977±2950.0 (100.0±0.0)	138209±2961.0 (100.0±0.0)

¹NFEV is defined as the average number of objective function evaluations during the optimization procedure and SR is the success rate in finding the global minimum based on 4 groups of 25 calculations with random initial values and number seeds.

²Global optimums are consistent with results reported by Gau et al. (2000).

Problem 4. Benzene-Hexafluorobenzene (EIV Formulation)

In this problem, we consider the modeling of vapor-liquid equilibrium of the binary system benzene-hexafluorobenzene using the error-in-variable approach. This system was studied by Gau and Stadtherr (2002) using interval analysis. Four data sets are used and parameter estimation is considered using the Wilson equation for liquid-phase activity coefficients. The state variables are (x_1, y_1, P, T) where P is the system pressure in mmHg, T is the system temperature in °C and x_1 and y_1 are the liquid and vapor mole fractions of benzene. In accordance with Gau and Stadtherr (2002), a standard deviations of 0.001, 0.01, 0.75 and 0.1 are assumed for these variables. The vapor-liquid equilibrium can be described by the following equations

$$P = \gamma_1 x_1 P_1^0 + \gamma_2 (1 - x_1) P_2^0 \quad (12)$$

$$y_1 = \frac{\gamma_1 x_1 P_1^0}{\gamma_1 x_1 P_1^0 + \gamma_2 (1 - x_1) P_2^0} \quad (13)$$

We formulate the data-fitting problem as unconstrained optimization using Eqs. [12] and [13]

to eliminate P and y_1 from the objective function. For the unconstrained problem, the independent state variables are the set of $\vec{z} = \begin{pmatrix} \vec{x}_1 \\ T \end{pmatrix}$ for all measurements, while the optimization variables are $\vec{\theta} = (\theta_1, \theta_2)^T$ and the set of $\vec{z}^t = \begin{pmatrix} \vec{x}_1^t \\ T^t \end{pmatrix}$. The objective function is defined as

$$F_{obj} = \sum_{i=1}^{ndat} \left[\frac{(x_{1i}^t - x_{1i})^2}{\sigma_{x_1}^2} + \frac{(y_{1i}^t - y_{1i})^2}{\sigma_{y_1}^2} + \frac{(T_i^t - T_i)^2}{\sigma_T^2} + \frac{(P_i^t - P_i)^2}{\sigma_P^2} \right] \quad (14)$$

This function is optimized with respect to $2 + 2ndat$ variables where $ndat$ is the number of experimental data used in parameter estimation. The initial intervals of the independent state variables are set using plus and minus three standard deviations, while the intervals for the Wilson model parameters are defined as $\theta_1 \in (-8500, 320000)$ and $\theta_2 \in (-8500, 320000)$. The results of parameter estimation are reported in Table 5 and the statistical

analysis appears in Table 2. Due to the increase in problem dimensionality, SA fails several times in finding the global minimum for all data sets. For only two cases, it shows 100% reliability in the global minimization of the objective function. Also, there is a significant increase in the computational effort of the data-fitting procedure, also caused by problem dimensionality. Of course, we can reduce the numerical effort of the SA method by decreasing the tolerance

value and using a local search method to refine the solution obtained by SA. On the other hand, the variance analysis indicates that parameter T_{SA} affects the reliability of parameter estimation for all data sets. For data at 760 mmHg, NT has a significant effect on the success rate for finding the global optimum. An increase in both T_{SA} and NT improves the reliability of SA. The global optimums reported for all data are consistent with the results of Gau and Stadtherr (2002).

Table 5: Results of parameter estimation for the benzene-hexafluorobenzene system using simulated annealing with the error-in-variable formulation

T, °C or P, mm Hg	Number of data	Global minimum ²			NT	NFEV(SR, %) ¹		
		θ_1	θ_2	F_{obj}		T_{SA}		
						10	100	1000
T = 50	11	-460.6	1115.8	19.525	10	756481±8521.0 (45.0±3.8)	826609±8081 (52.0±12.7)	896209±7906 (49.0±11.9)
					20	1510177±15659.0 (46.0±5.2)	1652353±13867 (53.0±11.9)	1787041±13602.0 (69.0±8.3)
70	9	-424.2	1006.8	8.503	10	867401±326470 (58.0±10.1)	965961±347410 (55.0±7.6)	994281±401374 (62.0±10.6)
					20	1534081±361615 (60.0±8.6)	1756321±616201 (58.0±10.6)	1645041±349990 (78.0±10.6)
P = 300	17	-478.3	1118.7	37.399	10	1157689±18171 (54.0±5.2)	1259281±19591 (56.0±5.7)	1360297±19289 (59.0±6.0)
					20	2296225±19676 (55.0±5.0)	2502865±17570 (56.0±5.7)	2709073±20534 (66.0±6.9)
760	29	-420.7	1060.3	16.925	10	4998841±3331470 (43.0±8.3)	4512481±2949822 (52.0±19.9)	4312441±3148616 (65.0±7.6)
					20	4105201±1095049 (96.0±3.3)	4236001±32451 (100.0±0.0)	4567861±32340 (100.0±0.0)

¹NFEV is defined as the average number of objective function evaluations during the optimization procedure and SR is the success rate in finding the global minimum based on 4 groups of 25 calculations with random initial values and number seeds.

²Global optimums are consistent with results reported by Gau and Stadtherr (2002).

Problem 5. Methanol-1,2 Dichloroethane

This system was studied by Esposito and Floudas (1998) using a deterministic global optimization method. We consider the EIV parameter estimation using the van Laar model for liquid phase. The experimental data consists of five points for the state variables: pressure in mmHg, temperature in Kelvin and liquid and vapor mole fractions of methanol. The vapor pressure of pure components is calculated using

$$P_i^0 = \exp\left(a_1 - \frac{b_1}{T - c_1}\right) \quad (15)$$

where P_i^0 is given in mmHg, T is given in Kelvin and the constants of the Antoine equation are taken

from Esposito and Floudas (1998). The activity coefficients are defined by

$$\gamma_1 = \exp\left[\frac{A}{RT} \left(1 + \frac{A}{B} \frac{x_1}{x_2}\right)^{-2}\right] \quad (16)$$

$$\gamma_2 = \exp\left[\frac{B}{RT} \left(1 + \frac{B}{A} \frac{x_2}{x_1}\right)^{-2}\right] \quad (17)$$

We use an unconstrained optimization formulation in the data-fitting procedure. In accordance with Esposito and Floudas (1998) and Gau and Stadtherr (2002), the temperature is scaled by a reference temperature $T_r = 323.15$ K and the van Laar model parameters are defined considering the reference

temperature as $\vec{\theta} = \left(\frac{A}{RT_r}, \frac{B}{RT_r} \right)$. The state variables of the system are $\left(x_1, y_1, P, \frac{T}{T_r} \right)$ and their standard deviations are defined as 0.005, 0.015, 0.75 and 0.000309. The objective function is given by

$$F_{\text{obj}} = \sum_{i=1}^{\text{ndat}} \left[\frac{(x_{1i}^t - x_{1i})^2}{\sigma_{x_1}^2} + \frac{(y_{1i}^t - y_{1i})^2}{\sigma_{y_1}^2} + \frac{(P_i^t - P_i)^2}{\sigma_P^2} + \frac{(TS_i^t - TS_i)^2}{\sigma_{TS}^2} \right] \quad (18)$$

where $TS = T/T_r$. Equations [12] and [13] are used to eliminate P and y_1 in the objective function and we have twelve optimization variables where the initial intervals for the independent state variables $\vec{z}^t = \left(\vec{x}_1^t, T^t \right)$ are set using plus and minus three standard deviations. For the case of the van Laar

model parameters, we use $\frac{A}{RT_r} \in (1, 2)$ and

$\frac{B}{RT_r} \in (1, 2)$. The results of parameter estimation are reported in Table 6. For all calculations performed, SA finds the global minimum of the objective function with 100% reliability. In this case, the parameters T_{SA} and NT affect only the efficiency of the data-fitting procedure. The global minimum reported is consistent with the results of Gau and Stadtherr (2002).

We have studied the numerical performance of SA with other examples reported by Gau et al. (2000) and Gau and Stadtherr (2002). The results obtained in those calculations, not reported here, indicate that SA is generally a reliable method for parameter estimation in VLE modeling.

Based on our experience with other SA algorithms, such as very fast simulated annealing (Sharma and Kaikkonen, 1999) and direct search simulated annealing (Ali et al., 2002), we can expect that the algorithm of Corana et al. (1987) has a better numerical performance and we consider it a suitable choice for data fitting in thermodynamic models.

Table 6: Results of parameter estimation for the methanol-1,2 dichloroethane system using simulated annealing with the error-in-variable formulation

T, °C	Number of data	Global minimum ²			NT	NFEV(SR, %) ¹		
		$\frac{A}{RT_r}$	$\frac{B}{RT_r}$	F_{obj}		T_{SA}		
						10	100	1000
50	5	1.912	1.608	3.326	10	370561±4435 (100.0±0.0)	404593±4766 (100.0±0.0)	438529±4370 (100.0±0.0)
					20	742273±9166 (100.0±0.0)	810433±8783 (100.0±0.0)	878689±7863 (100.0±0.0)

¹NFEV is defined as the average number of objective function evaluations during the optimization procedure and SR is the success rate in finding the global minimum based on 4 groups of 25 calculations with random initial values and number seeds.

²Global optimum is consistent with results reported by Gau and Stadtherr (2002).

CONCLUSIONS

The numerical performance of the simulated annealing method has been tested in parameter estimation for VLE modeling. The reliability and efficiency of the data-fitting procedure was evaluated with respect to the SA parameters T_{SA} and NT , using classical least squares and error-in-variable formulations. Our results indicate that parameter T_{SA} significantly affects the reliability of the data-fitting procedure. On the other hand, both parameters affect the efficiency of the SA method. The SA method has a better numerical performance using the classical

least squares instead of the error-in-variables formulation. In general, SA is a robust optimization procedure, when properly implemented for parameter estimation using both formulations. However, it can fail in the global minimization of difficult systems.

ACKNOWLEDGMENTS

The authors acknowledge the financial support received from CONACyT and Instituto Tecnológico de Aguascalientes.

NOMENCLATURE

A, B	parameters of the van Laar model
a_1, b_1, c_1	parameters of the Antoine equation
NT	iteration number before temperature reduction
P	Pressure
P_i^0	vapor pressure of pure component
T	temperature
T_{SA}	annealing temperature
T_r	reference temperature
x_1, y_1	liquid and vapor mole fraction
y_{ij}	dependent response variable
z_{ij}^t	unknown "true" value of state variable
σ_i	standard deviation of state variable
θ_1, θ_2	energy parameters
γ	activity coefficient

REFERENCES

- Ali, M.M., Törn, A., and Viitanen, S., A direct search variant of the simulated annealing algorithm for optimization involving continuous variables. *Computers & Operations Research*, 29 (1), 87-102 (2002).
- Anderson, T.F., Abrams, D.S., and Grens, E.A. II., Evaluation of parameters for nonlinear thermodynamics models. *AIChE Journal*, 24 (1), 20-29 (1978).
- Bard, Y., *Nonlinear parameter estimation*, Academic Press, New York (1974).
- Britt, H.I., and Luecke, R.H., The estimation of parameters in nonlinear implicit models. *Technometrics*, 15, 233-247 (1973).
- Corana, A., Marchesi, M., Martini, C., and Ridella, S., Minimizing multimodal functions of continuous variables with the simulated annealing algorithm. *ACM Transactions on Mathematical Software*, 13 (3), 262-280 (1987).
- Costa, A.L.H., da Silva, F.P.T., and Pessoa, F.L.P., Parameter estimation of thermodynamic models for high-pressure systems employing a stochastic method of global optimization. *Brazilian Journal of Chemical Engineering*, 17 (3), 349-353 (2000).
- Dominguez, A., Tojo, J., and Castier, M., Automatic implementation of thermodynamic models for reliable parameter estimation using computer algebra. *Computers and Chemical Engineering*, 26 (10), 1473-1479 (2002).
- Englezos, P., and Kalogerakis, N., *Applied parameter estimation for chemical engineers*, Marcel-Dekker, New York (2001).
- Esposito, W.R., and Floudas, C.A., Global optimization in parameter estimation of nonlinear algebraic models via the error-in-variables approach. *Ing. Eng. Chem. Res.*, 37 (5), 1841-1858 (1998).
- Fabries, J., and Renon, H., Method of evaluation and reduction of vapor-liquid equilibrium data of binary mixtures. *AIChE Journal*, 21 (4), 735-743 (1975).
- Gau C.Y., Brennecke J.F., Stadtherr M.A., Reliable nonlinear parameter estimation in VLE modeling. *Fluid Phase Equilibria*, 168 (1), 1-18 (2000).
- Gau, C.Y., and Stadtherr, M.A., Deterministic global optimization for error-in-variables parameter estimation. *AIChE Journal*, 48 (6), 1192-1197 (2002).
- Gmehling, J., Onken, U., and Arlt, W., *Vapor-liquid Equilibrium Data Collection*, Chemistry Data Series, Vol. I, Parts 1-8, DECHEMA (1977-1990).
- Goffe, B., Ferrier, G., and Rogers, J., Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, 60 (1-2), 65-99 (1994).
- Henderson, N., de Oliveria, J.R., Amaral Souto, H.P., and Pitanga Marques, R., Modeling and analysis of the isothermal flash problem and its calculation with the simulated annealing algorithm. *Ind. Eng. Chem. Res.*, 40 (25), 6028-6038 (2001).
- Kirkpatrick, S., Gelatt, C., and Vecchi, M.P., Optimization by simulated annealing. *Science*, 220 (4598), 671-680 (1983).
- Kleiber, M., and Axmann, J.K., Evolutionary algorithms for the optimization of modified UNIFAC parameters. *Computers and Chemical Engineering*, 23 (1), 63-82 (1998).
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller A., and Teller, E., Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21 (6), 1087-1092 (1953).
- Michalewicz Z., Fogel D.B., *How to solve it? Modern heuristics*, Springer (1999).
- Park, T.Y., and Froment, G.F., A hybrid genetic algorithm for the estimation of parameters in detailed kinetic models. *Computers and Chemical Engineering*, 22 (S1), S103-S110 (1998).
- Rangaiah, G.P., Evaluation of genetic algorithms and simulated annealing for phase equilibrium and stability problems. *Fluid Phase Equilibria*, 187-188, 83-109 (2001).

- Reid, R.C., Prausnitz, J.M., and Sherwood, T.K., The properties of gases and liquids, 3rd Ed., McGraw-Hill, New York (1987).
- Sharma, S.P., and Kaikkonen, P., Global optimisation of time domain electromagnetic data using very fast simulated annealing. *Pure and Applied Geophysics*, 155 (1), 149-168 (1999).
- Tjoa, T.B., and Biegler, L.T., Reduced successive quadratic programming strategy for errors-in-variables estimation. *Computers Chemical Engineering*, 16 (6), 523-533 (1992).
- Valko, P., and Vajda, S., An extended marquardt-type procedure for fitting error-in-variables methods. *Computers Chemical Engineering*, 11 (1), 37-43 (1987).
- Vamos, R.J., and Hass C.N., Reduction of ion-exchange equilibria data using an error in variable approach. *AIChE J.*, 40 (3), 556 – 569 (1994).