

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

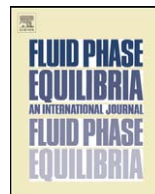
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Fluid Phase Equilibria

journal homepage: www.elsevier.com/locate/fluid

Evaluation of stochastic global optimization methods for modeling vapor–liquid equilibrium data

Adrián Bonilla-Petriciolet^a, Gade Pandu Rangaiah^{b,*}, Juan Gabriel Segovia-Hernández^c

^a Department of Chemical Engineering, Instituto Tecnológico de Aguascalientes, Aguascalientes, 20256, Mexico

^b Department of Chemical & Biomolecular Engineering, National University of Singapore, Singapore, 117576, Singapore

^c Department of Chemical Engineering, Universidad de Guanajuato, Campus Guanajuato, Guanajuato, 36050, Mexico

ARTICLE INFO

Article history:

Received 9 July 2009

Received in revised form

16 September 2009

Accepted 22 September 2009

Available online 27 September 2009

Keywords:

Vapor–liquid equilibrium

Simulated Annealing

Genetic Algorithm

Differential Evolution

Particle Swarm Optimization

ABSTRACT

Parameter estimation for vapor–liquid equilibrium (VLE) data modeling plays an important role in design, optimization and control of separation units. This optimization problem is very challenging due to the high non-linearity of thermodynamic models. Recently, several stochastic optimization methods such as Differential Evolution with Tabu List (DETL) and Particle Swarm Optimization (PSO) have evolved as alternative and reliable strategies for solving global optimization problems including parameter estimation in thermodynamic models. However, these methods have not been applied and compared with respect to other stochastic strategies such as Simulated Annealing (SA), Differential Evolution (DE) and Genetic Algorithm (GA) in the context of parameter estimation for VLE data modeling. Therefore, in this study several stochastic optimization methods are applied to solve parameter estimation problems for VLE modeling using both the classical least squares and maximum likelihood approaches. Specifically, we have tested and compared the reliability and efficiency of SA, GA, DE, DETL and PSO for modeling several binary VLE data using local composition models. These methods were also tested on benchmark problems for global optimization. Our results show that the effectiveness of these stochastic methods varies significantly between the different tested problems and also depends on the stopping criterion especially for SA, GA and PSO. Overall, DE and DETL have better performance for solving the parameter estimation problems in VLE data modeling.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The estimation of parameters in thermodynamic models is an important requirement and a common task in many areas of chemical engineering because these models form the basis for synthesis, design, optimization and control of process systems [1,2]. In the particular case of process separation, thermodynamic models play a major role with respect to energy requirements, phase equilibria and equipment sizing. Specifically, the parameter estimation problem refers to determining values of the model parameters that provide the best fit to a set of measured data [3,4].

Data modeling using thermodynamic equations is generally based on classical least squares or maximum likelihood approaches [1,3,5–7]. In the classical least squares, it is assumed that independent variables are not subject to measurement error, while errors in all measured variables are accounted in the maximum likelihood approach. Both formulations involve the minimization of a suitable objective function subject to constraints arising from the

model equations. Data correlation can be performed either by direct optimization of the objective function or by solving an equivalent system of non-linear equations obtained from the stationary conditions of the optimization problem. Most of thermodynamic models are non-linear in their adjustable parameters and, as a consequence, the objective function for data fitting is non-linear and potentially non-convex. Therefore, parameter estimation problems are often very difficult to solve reliably even for simple mathematical equations [1,3].

In particular, estimation of parameters in non-linear thermodynamic models for vapor–liquid equilibrium (VLE) modeling has been of great interest in the chemical engineering literature. Experimental VLE data are essential for the design of industrial separation processes and for development of accurate models in process systems engineering [8]. However, a number of pitfalls and difficulties may be faced in parameter estimation for VLE modeling; these include: convergence to a local minimum, flat objective function in the neighborhood of the globally optimal solution, badly scaled model functions, and non-differentiable terms in thermodynamic equations. In addition, the number of optimization variables can be very large especially for error-in-variable formulations. Several researchers have demonstrated the challenging nature of param-

* Corresponding author. Tel.: +65 65162187; fax: +65 6779 1936.

E-mail address: chepr@nus.edu.sg (G.P. Rangaiah).

eter estimation problems for VLE data modeling, and they have highlighted the need for reliable numerical techniques in order to overcome these difficulties, e.g., [1,3,4,6,7,9–13].

Note that failure to find the globally optimal parameters for a thermodynamic model and using locally optimal parameters instead, can have significant consequences in subsequent calculations, may cause errors and uncertainties in equipment design and erroneous conclusions about model performance [3]. In the context of VLE data modeling, recent studies have shown that using the locally optimal parameters may result in incorrect predictions of the azeotropic states with local composition models and in qualitative discrepancies of the phase behavior such as prediction of spurious phase split and modeling of homogeneous azeotropes as heterogeneous [3,13]. These failures are undoubtedly potential sources of problems for design of separation processes.

For reliable parameter estimation in VLE modeling, the use of deterministic and stochastic global optimization methods has been suggested in the literature. The first type of methods guarantees convergence to the global solution, and they include two main approaches: branch-and-bound optimization procedures with convex underestimating functions [1] and interval techniques [3,6,7,9]. Recently, a bi-level optimization with non-convex lower-level programs has been proposed for reliable parameter estimation in VLE problems [13]. Unfortunately, current deterministic global optimization methods may be too expensive (in terms of computational effort) for multivariable problems and, in some cases, problem reformulation is needed depending on the characteristics of the thermodynamic model under study. In contrast, stochastic global optimization methods can locate good solutions in a reasonable CPU time, are easy to use and independent of model although they do not guarantee the global optimality. These methods usually converge quickly to the vicinity of the global solution and are suitable to deal with large number of parameters. It appears that these methods may offer the best compromise between quality of solution and efficiency for multivariable parameter estimation problems.

Until now, Simulated Annealing (SA), Genetic Algorithm (GA), Random Tunneling algorithm (RTA) and Differential Evolution (DE) have been successfully used for solving parameter estimation in VLE modeling [4,10–12,14,15]. Specifically, Alvarez et al. [4] applied and compared two versions of GA for VLE modeling using local composition models and equations of state. Srinivas and Rangiah [10] used the RTA for VLE modeling using the error-in-variable (EIV) approach. Bonilla-Petriciolet et al. [11] studied the performance of SA for parameter estimation in VLE modeling using both least squares and maximum likelihood formulations. In another study, DE was successfully applied to modeling the equilibrium solubility of CO₂ in aqueous alkanolamines [12]. Costa et al. [14] reported the application of SA for parameter estimation in the modeling of vapor–solid equilibrium with supercritical carbon dioxide as the solvent. Steyer and Sundmacher [15] used an evolutionary optimization strategy for the simultaneous fitting of VLE and liquid–liquid equilibrium (LLE) data for ternary systems. Apart from the methods used in these studies, other reliable stochastic methods such as Differential Evolution with Tabu List (DETL) and Particle Swarm Optimization (PSO) are available in the literature and have been successfully used for thermodynamic calculations, e.g. [16,17]; so, they are promising for parameter estimation in thermodynamic models. To the best of the authors' knowledge, these novel methods have not been tested and evaluated for VLE data modeling.

In this study, we apply and compare a number of stochastic global optimization methods for modeling VLE data. Specifically, we study the performance of SA, GA, DE, PSO and DETL for VLE data modeling of several binary systems with local composition models using the classical least squares and maximum likelihood approaches, as well as for benchmark problems. This comparison

is essential to identify the relative strengths of novel stochastic optimization methods for VLE data modeling. This study is the first attempt to compare comprehensively a number of stochastic global optimization methods for parameter estimation in VLE data modeling.

2. Description of stochastic optimization methods

Many stochastic global methods have been proposed and tested in several disciplines, and they are expected to be widely used in chemical engineering including thermodynamic applications in the coming years. In this study, we used five stochastic optimization methods: SA, GA, DE, DETL and PSO for parameter estimation in VLE data modeling. Note that SA, GA and DE are classical stochastic optimization methods and have recently been used by the authors and other researchers, while DETL and PSO are new to parameter estimation problems in VLE modeling. Further, PSO has been selected instead of Ant Colony Optimization or other swarm intelligence methods due to the results reported in other global optimization studies, e.g. [18], which indicate that PSO may perform better. All the selected methods have the attributes of a good optimization strategy such as generality, efficiency, reliability and ease of use [19]. SA is a point-to-point method while GA, DE, DETL and PSO are population-based methods. A brief description of these algorithms is presented in the following subsections, and detailed explanation of them is available in the cited references.

In this paper, the global optimization problem for VLE modeling is as follows: minimize $F_{obj}(\vec{u})$ subject to $\vec{u} \in \Omega$ where \vec{u} is a continuous variable vector in the domain Ω in n_{var} dimensions, and $F_{obj}(\vec{u}): \Omega \rightarrow \Re$ is a real-valued function. The minimization of F_{obj} can be treated either as a constrained or, by substitution of the model equations in the objective function, as an unconstrained minimization problem. In the present study, we have solved the unconstrained optimization problem with the domain Ω defined by the upper and lower limits of each decision variable. The n_{var} decision variables in the VLE parameter estimation problems are the adjustable parameters of thermodynamic models if least squares criterion is used. When the EIV approach is applied, the decision variables are both these model parameters and the true values of state variables (i.e. temperature, pressure, and liquid and vapor compositions).

2.1. Simulated Annealing

SA mimics the thermodynamic process of cooling of molten metals to attain the lowest free energy state [20]. Starting with an initial solution, the algorithm performs a stochastic partial search of the space defined for decision variables. In minimization problems, uphill moves are occasionally accepted with a probability controlled by the parameter called annealing temperature: T_{SA} . The probability of acceptance of uphill moves decreases as T_{SA} decreases. At high T_{SA} , the search is almost random, while at low T_{SA} the search becomes selective where good moves are favored. The core of this algorithm is the Metropolis criterion [21], used to accept or reject uphill movements with the acceptance probability given by

$$M(T_{SA}) = \min \left\{ 1, \exp \left(\frac{-\Delta f}{T_{SA}} \right) \right\} \quad (1)$$

where Δf is the change in objective function value from the current point to new point.

In this study, the SA algorithm proposed by Corana et al. [22] has been used because of its good performance in thermodynamic calculations, e.g. [23,24]. In this algorithm, a trial point is randomly chosen within the step length VM (which is a vector of length n_{var}) from the starting/current point. The objective function is evalu-

ated at this trial point, and its value is compared to the objective value at the starting/current point. Eq. (1) is used to accept or reject the trial point. If this trial point is accepted, the algorithm continues the search using that point; otherwise, another trial point is generated within the neighborhood of the starting/current point. 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 $NT \times NS \times n_{var}$ function evaluations, is imposed upon the system using the cooling schedule. Note that NT is the number of iterations before T_{SA} reduction and NS is the number of cycles for updating the decision variables.

In our calculations, cooling schedule for decreasing T_{SA} is defined as

$$T_{SA,k} = 0.5(T_{SA,0} - T_{SA,F}) \left(1 - \tanh \left(\frac{17k}{Iter_{max}} - 5 \right) \right) + T_{SA,F} \quad (2)$$

where $Iter_{max}$ is the maximum number of iterations for SA, $T_{SA,k}$ is the annealing temperature at iteration k , and $T_{SA,0}$ and $T_{SA,F}$ are the initial and final values for the annealing temperature, respectively. Thus, as T_{SA} declines, downhill moves are less likely to be accepted and SA focuses on the most promising area for optimization. The iterative steps are performed until the specified stopping criterion: either the maximum number of successive iterations (Sc_{max}) without improvement in the best function value, or until the maximum number of iterations ($Iter_{max}$), is satisfied. The main parameters of SA are $T_{SA,0}$, $T_{SA,F}$, NS , NT , Sc_{max} and $Iter_{max}$. Detailed description and flowchart of this algorithm can be found in Corana et al. [22]. We have used, after suitable modifications, the subroutine developed by Goffe et al. [25], for the present study.

2.2. Genetic Algorithm

GA is a stochastic technique that simulates natural evolution on the solution space of the optimization problems. It operates on a population of potential solutions (individuals) in each iteration (generation). By combining some individuals of the current population according to predefined operations, a new population that contains better individuals, is produced as the next generation. The first step of GA is to create randomly an initial population of N_{pop} solutions in the feasible region. GA works on this population and combines (crossover) and modifies (mutation) some chromosomes according to specified genetic operations, to generate a new population with better characteristics. Individuals for reproduction are selected based on their objective function values and the Darwinian principle of the survival of the fittest [26]. Genetic operators are used to create new individuals for the next population from those selected individuals of the current population, and they serve as searching mechanisms in GA. In particular, crossover forms two new individuals by first choosing two individuals from the mating pool (containing the selected individuals) and then swapping different parts of genetic information between them. This combining (crossover) operation takes place with a user-defined crossover probability (P_{cros}) so that some parents remain unchanged even if they are chosen for reproduction. Mutation is an unary operator that creates a new solution by a random change in an individual. It provides a guarantee that the probability of searching any given string will never be zero and acting as a safety net to recover good genetic material which may be lost through the action of selection and crossover. The mutation procedure proceeds with a probability P_{mut} .

Selection, crossover and mutation procedures are recursively used to improve the population and to identify promising areas for optimization. This algorithm terminates when the user-specified criterion is satisfied. For comparison purposes, the stopping conditions described for SA have been implemented in all stochastic methods tested in this work. Specifically, GA stops after evolving

for the specified number of generations (Gen_{max}), or until performing the maximum number of successive generations (Sc_{max}) without improvement in the best objective value. We have used a GA with floating-point encoding, selection via stochastic universal sampling, modified arithmetic crossover and non-uniform mutation. Details of this algorithm are available in Rangaiah [23]. The key parameters of GA are N_{pop} , P_{cros} , P_{mut} , Gen_{max} and Sc_{max} .

2.3. Particle Swarm Optimization

PSO is a novel and promising population-based method that belongs to the class of swarm intelligence algorithms. Kennedy and Eberhart [27] introduced this strategy for global optimization, which is inspired by the social behavior of flocking swarms of birds and fish schools. It exploits a population of potential solutions to identify promising areas for optimization. In this context, the population of potential solutions is called the *swarm* and each solution is called *particle*. *Particles* are conceptual entities, which fly through the multi-dimensional search space. The success histories of the particles influence both their own search patterns and those of their peers. Each particle has two state variables: its current position and current velocity. In the PSO version used in this study, the search is focused on promising regions by biasing each particle's velocity towards both the particle's own remembered best position and the communicated best neighborhood location so far. The relative weights of these two positions are scaled by the cognitive and social parameters. Cognitive parameter has a contribution towards the self-exploration (or experience) of a particle, while social parameter has a contribution towards motion of the particles in global direction taking into account the swarm motion in the preceding iteration [28].

Thus, the velocity and position of each particle is updated using

$$V_{i,j}(k+1) = wV_{i,j}(k) + C_1 R_1 (s_{i,j}^p - s_{i,j}(k)) + C_2 R_2 (s_{i,j}^{best} - s_{i,j}(k)) \quad (3)$$

$$s_{i,j}(k+1) = s_{i,j}(k) + V_{i,j}(k+1) \quad (4)$$

where $R_1, R_2 \in (0, 1)$ are random numbers, C_1 and C_2 are the cognitive and social parameters, $s_{i,j}(k)$ is the position/value of decision variable i in particle j at iteration k , $V_{i,j}(k)$ is the velocity of decision variable i in particle j at iteration k , $s_{i,j}^p$ is the particle's own remembered best position for decision variable i in particle j , $s_{i,j}^{best}$ is the communicated best neighborhood location so far for decision variable i in particle j , and w is the inertia weight factor.

The inertia weight factor is used to control the impact of the previous velocities on the current velocity. It influences the trade-off between the global and local exploration abilities of the particles. At the initial stage of the search process, large inertia weight is recommended to enhance the global exploration; at the last stage, this parameter should be reduced for better local exploration. Therefore, in our calculations, we have considered w decreasing linearly from w_0 to w_f over the whole run; and C_1 also decreases linearly from $C_{1,0}$ to $C_{1,f}$. Specifically, we have used the following equations:

$$C_1 = (C_{1,f} - C_{1,0}) \left(\frac{k}{Iter_{max}} \right) + C_{1,0} \quad (5)$$

$$w = (w_f - w_0) \left(\frac{k}{Iter_{max}} \right) + w_0 \quad (6)$$

where $Iter_{max}$ is the maximum number of iterations for PSO. Previous numerical experience indicates that the dynamic adaptation of these parameters improves the global search over the entire search space during the early iterations and leads the particles to converge to global optimum at the end of the search [28]. Usually C_1 ranges from [0.5, 3.5] whereas w ranges [0.4, 0.8]. Based on preliminary trials and results reported in the literature, we have chosen $C_{1,f} = 0.5$, $w_f = 0.4$ and $C_2 = 4.0 - C_1$ in the present study. In addition,

tion, the velocity of each particle is restricted to a maximum value within the interval $[-V_{\max}, V_{\max}]$, which is defined considering the bounds on decision variables. Each particle is assigned to a neighborhood of a pre-specified number of particles, n_h ($=0.25 \times$ number of particles, n_p). The best position attained so far by particles that comprise the neighborhood is communicated among them.

The overall algorithm of PSO is as follows. At the beginning, a population of n_p particles is initialized with random positions s_{ij} and random velocities V_{ij} for $i=1, 2, \dots, n_{\text{var}}$ and $j=1, 2, \dots, n_p$. Once the n_p particles are initialized, the positions and velocities of all particles are modified using Eqs. (3) and (4). After calculating the velocities and position for the next iteration $k+1$, the current iteration is completed. The best particle is updated only when a new one is found yielding a decrease in the objective function value. This process is performed for a certain number of iterations ($Iter_{\max}$), or until the maximum number of successive iterations (Sc_{\max}) without improvement in the best function value has been reached. The main parameters of PSO are: n_p , w_0 , $C_{1,0}$, $Iter_{\max}$ and Sc_{\max} .

2.4. Differential Evolution

DE is a population-based method that has found several applications in science and engineering, including chemical engineering and for parameter estimation. The algorithm used in this study starts with specifying the parameters, namely: amplification factor (A), crossover constant (CR), population size (NP), maximum number of successive iterations (Sc_{\max}) without improvement in the best function value and maximum number of generations (Gen_{\max}). The initial population is randomly generated using the uniformly distributed random numbers to cover the entire feasible space. The individuals are checked for boundary violation to see if any individual is generated in the infeasible region; the infeasible points are replaced by generating new individuals. The objective function values of all the individuals are calculated, and the best point is determined.

Then, the three main steps: mutation, crossover and selection on the population, are carried out. Mutation and crossover operations are performed to diversify the search thus escaping from the local minima. The mutant vector is generated for each randomly chosen target vector $X_{i,G}$ by

$$V_{i,G+1} = X_{R1,G} + A(X_{R2,G} - X_{R3,G}) \quad i = 1, 2, 3, \dots, NP \quad (7)$$

where random numbers R_1 , R_2 and R_3 are distinct and belong to the set $\{1, 2, \dots, NP\}$, and $X_{R1,G}$, $X_{R2,G}$ and $X_{R3,G}$ represents the three random individuals chosen in the current generation G , to produce the mutant vector for the next generation, $V_{i,G+1}$. The random numbers should be different from the running index, i , and hence NP should be ≥ 4 to allow mutation. Parameter A is a real value between 0 and 2, and it controls the amplification of the differential variation between the two random individuals.

In the crossover step, the trial vector $U_{i,G+1}$ is produced by copying some elements of the mutant vector, $V_{i,G+1}$ to the target vector, $X_{i,G}$ with probability equal to CR . A random number (ran) is generated for each element of the target vector; if $ran \leq CR$, the element of the mutant vector is copied to the trial vector, else the target vector element is copied. After mutation and crossover operations, the trial vector competes with the target vector for selection into the next generation. A greedy criterion based on the objective value is used for this selection. If the trial vector has a better value compared to the target vector, it replaces the target vector in the population thus allowing the better solution into further generations. The process of mutation, crossover and selection is repeated until the termination criterion (Gen_{\max} or Sc_{\max}) is satisfied. The algorithm then terminates providing the best point that has been explored over all the generations. A detailed explanation of this method is provided by Storn and Price [29], and we have used the

algorithm reported by Srinivas and Rangaiah [30] in the present study.

2.5. Differential Evolution with Tabu List

This recent stochastic method developed by Srinivas and Rangaiah [16] is a hybrid strategy obtained from DE and Tabu Search (TS). DETL begins with the selection of values for parameters: population size (NP), amplification factor (A), crossover constant (CR), tabu radius (tr), tabu list size (tls), maximum number of generations (Gen_{\max}), and maximum number of successive generations (Sc_{\max}) without improvement in the best function value. The algorithm generates the initial population of size NP using uniformly distributed random numbers to cover the entire feasible region. The objective function is evaluated at each individual/point, and the best one is selected. The tabu concept of TS is implemented in the generation step of DE (i.e., after crossover and mutation) to improve the diversity among the individuals and consequently the computational efficiency. It employs a tabu list with the parameters: tr and tls , which keep track of the evaluated points for avoiding revisits to them during the subsequent search.

The three main steps: mutation, crossover and selection of DE along with tabu checking are performed on the population during each generation. For this, a mutant individual is generated for each randomly chosen target individual ($X_{i,j}$) in the population by Eq. (7). In the crossover step, a trial individual/vector is generated by copying some elements of the mutant individual to the target individual with a probability of CR . A boundary violation check is performed to check the feasibility of the resulting trial individual; if any bound is violated, the trial individual is either replaced by generating a new individual or forced to the nearest boundary (lower or upper). The trial individual is then compared to the already evaluated points in the tabu list in terms of the Euclidean distance. If the Euclidean distance is smaller than the tabu radius, which indicates that the objective function value at the trial vector and at one of the points in the tabu list are probably close to each other, the trial individual is rejected considering that it may not give new information about the objective function except increasing the number of function evaluations. The rejected point is replaced by generating another trial point by crossover and mutation operations, until the Euclidean distance between the new point and to all points in the tabu list is greater than the tabu radius.

The process of generating new individuals, including checking their closeness to those in the tabu list, is repeated until all members of the new population are produced. The objective function is evaluated at the trial individual only if it is away from all the points in the tabu list. After each evaluation, the tabu list is updated dynamically to keep the latest point(s) in the list by replacing the earliest entered point(s). In the selection step, a greedy criterion such as fitness (i.e., objective function) value is used to select the better one between the trial and target individuals. If the trial individual is selected, it replaces the target individual in the population immediately and may participate in the subsequent mutation and crossover operations. If the target individual is better, then it remains in the population and may participate in the subsequent mutation and crossover operations. The process of generation, evaluation and selection is repeated NP times in each generation. The algorithm runs until the stopping criterion (Gen_{\max} or Sc_{\max}) is satisfied, and gives the best point obtained over all the generations. Details of DETL algorithm can be found in Srinivas and Rangaiah [16].

2.6. Implementation of the methods

In the present study, FORTRAN codes developed for the five stochastic algorithms were used. These codes are available to interested readers upon request to the corresponding author. Each

method has been implemented in combination with a local optimization technique at the end of global search, for finding the global minimum accurately and efficiently. Quasi-Newton method implemented in the subroutine DBCONF of IMSL library was used for local optimization. This subroutine calculates the gradient via finite differences and approximates the Hessian matrix according to BFGS formula. For more details on this local strategy, see the optimization book by Dennis and Schnabel [31]. The default values of DBCONF parameters in the IMSL library were used in our study. All calculations were performed on the Intel Pentium M 1.73 GHz processor with 504 MB of RAM. This computer performs 254 million floating point operations per second (MFlops) for the LINPACK benchmark program (available at <http://www.netlib.org/>) for a matrix of order 500.

3. Formulation of parameter estimation problem for VLE modeling

Consider a set of observations q_{ij} of $i=1, 2, \dots, m$ dependent/response variables from $j=1, 2, \dots, ndat$ experiments are available, where the responses can be expressed by an explicit model $q_{ij} = f_i(\vec{r}_j, \vec{\theta})$, with nl independent variables $\vec{r}_j = (r_{1,j}, \dots, r_{nl,j})^T$ and $npar$ parameters $\vec{\theta} = (\theta_1, \dots, \theta_{npar})^T$. Measurement errors in \vec{r}_j can either be treated or neglected; depending on this choice, we can have a least squares or maximum likelihood formulation [2,3]. Only the model parameters are determined in the first approach whereas, in the second formulation, both the true values of state variables and model parameters are obtained [2]. We have used both formulations for modeling VLE data, and the performance of the selected stochastic methods has been compared in these scenarios.

Slightly different objective functions can be used to obtain the parameter values that provide the best fit for a specified model [32]. Objective functions involving fractional errors (see Eq. (8)) are preferred for data fitting using thermodynamic equations because they weigh errors in small and large quantities equally. Thus, for the case of classical least squares (LS) criterion, the following objective function is used.

$$F_{obj} = \sum_{j=1}^{ndat} \sum_{i=1}^m \left(\frac{q_{ij} - f_i(\vec{r}_j, \vec{\theta})}{q_{ij}} \right)^2 \quad (8)$$

where q_{ij} is the set of observations, f_i is the model for dependent variable i , $\vec{\theta}$ is the vector of parameters, \vec{r}_j is the vector of independent variables, m is the number of dependent/response variables and $ndat$ is the number of experiments. This function is minimized with respect to the model parameters $\vec{\theta}$ inside specified bounds. 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. Therefore, the objective function commonly used for data fitting is based on activity coefficients

$$F_{obj} = \sum_{j=1}^{ndat} \sum_{i=1}^c \left(\frac{\gamma_{ij}^{exp} - \gamma_{ij}^{calc}}{\gamma_{ij}^{exp}} \right)^2 \quad (9)$$

where γ_i^{exp} and γ_i^{calc} are respectively the experimental and calculated values for the activity coefficient of component i and c is the number of components in the mixture. Note that this formulation is equivalent to assuming that the standard error in the measurement of γ_{ij} is proportional to its value [2].

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, γ_i^{exp} can be calculated from VLE data using the following expression:

$$\gamma_i^{exp} = \frac{y_i^{exp} P}{x_i^{exp} P_i^0} \quad i = 1, \dots, c \quad (10)$$

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 . We have used the Wilson, NRTL and UNIQUAC models to calculate the liquid-phase activity coefficients, γ_i^{calc} , and Eq. (9) is optimized with respect to the energy parameters of these models. The energy parameters of thermodynamic models are defined as in the DECHEMA [33]. The global minimization of LS objective function can be done as an unconstrained optimization problem using local composition models.

On the other hand, if we assume that there are measurement errors in the state variables z_{ij} for the experiments of the system to be modeled, the minimization problem that must be solved is the error-in-variable (EIV) formulation of the form:

$$F_{obj} = \sum_{j=1}^{ndat} \sum_{i=1}^{nest} \frac{(z_{ij}^t - z_{ij})^2}{\sigma_i^2} \quad (11)$$

subject to

$$\vec{g}(z_{ij}^t, \vec{\theta}) = 0 \quad i = 1, \dots, nest \quad j = 1, \dots, ndat \quad (12)$$

where \vec{g} is a vector of np model functions, $nest$ is the number of state variables, z_{ij}^t is the unknown "true" value of i th state variable in j th experiment, and σ_i is the standard deviation associated with the measurement of i th state variable. The decision variables of EIV problem are the set of z_{ij}^t and the model parameters $\vec{\theta}$. In this formulation, there is a substantial increase in the dimensionality of the optimization problem, which depends on the number of experiments ($ndat$). For the case of VLE data, the state variables are x , y , P and T with standard deviations (σ_x , σ_y , σ_P and σ_T).

At low pressure, VLE can be described by the equations:

$$P = \sum_{i=1}^c \gamma_i x_i P_i^0 \quad (13)$$

$$y_i = \frac{\gamma_i x_i P_i^0}{\sum_{j=1}^c \gamma_j x_j P_j^0} \quad i = 1, \dots, c \quad (14)$$

where γ_i is the calculated activity coefficient by the chosen thermodynamic model and c is the number of components involved in the VLE system. Then, we can formulate the data fitting problem as an unconstrained optimization problem using Eqs. (13) and (14) to eliminate P^t and y_i^t in the EIV objective function (Eq. (15)). For the unconstrained problem, the independent variables are the set of $\vec{z} = (\vec{x}_{ij}, \vec{T})$ for all measurements, while the decision variables are $\vec{\theta} = (\theta_1, \dots, \theta_{npar})^T$ and the set of $\vec{z}^t = (\vec{x}_{ij}^t, \vec{T}^t)$. The objective function for VLE data correlation using the EIV approach can be defined as

$$F_{obj} = \sum_{j=1}^{ndat} \sum_{i=1}^c \left[\frac{(x_{ij}^t - x_{ij})^2}{\sigma_{x_i}^2} + \frac{(y_{ij}^t - y_{ij})^2}{\sigma_{y_i}^2} + \frac{(T_j^t - T_j)^2}{\sigma_T^2} + \frac{(P_j^t - P_j)^2}{\sigma_P^2} \right] \quad (15)$$

which is optimized with respect to $npar + c \times ndat$ decision variables. Note that mole fraction summation is one equality constraint, which was used to eliminate one decision variable (mole fraction), for each dataset.

For both LS and EIV formulations, the highly non-linear form of the thermodynamic models makes F_{obj} strongly non-linear, poten-

Table 1
Details of the benchmark problems used for performance comparison of stochastic optimization methods.

F_{obj}	Number of decision variables, n_{var}	Global minimum	Remarks
Zakharov (ZAK $_{n_{var}}$), $F_{obj} = \left(\sum_{i=1}^{n_{var}} u_i^2 \right) + \left(\sum_{i=1}^{n_{var}} 0.5iu_i^2 \right)^2 + \left(\sum_{i=1}^{n_{var}} 0.5iu_i^2 \right)^4$	$n_{var} = 2, 5, 10$ and 20 ; $-5 \leq u_i \leq 10$	0 at $u = (0, \dots, 0)$	Unimodal
Rosenbrock (ROS $_{n_{var}}$), $F_{obj} = \sum_{i=1}^{n_{var}} [100(u_i^2 - u_{i+1})^2 + (u_i - 1)^2]$	$n_{var} = 2, 5, 10$ and 20 ; $-5 \leq u_i \leq 10$	0 at $u = (1, \dots, 1)$	The global optimum is inside a long, narrow, parabolic shaped flat valley
Goldstein and Price (GP $_2$), $F_{obj} = [1 + (u_1 + u_2 + 1)^2(19 - 14u_1 + 3u_1^2 - 14u_2 + 6u_1u_2 + 3u_2^2)] * [30 + (2u_1 - 3u_2)^2(18 - 32u_1 + 12u_1^2 + 48u_2 - 36u_1u_2 + 27u_2^2)]$	$n_{var} = 2$; $-2 \leq u_i \leq 2$	3 at $u = (0, -1)$	4 local minima
Modified Himmelblau (mHB), $F_{obj} = (u_1^2 + u_2 - 11)^2 + (u_1 + u_2 - 7)^2 + 0.1((u_1 - 3)^2 + (u_2 - 2)^2)$	$n_{var} = 2$; $-6 \leq u_i \leq 6$	0 at $u = (3, 2)$	4 local minima
Rastrigin (RAS $_{n_{var}}$), $F_{obj} = 10n_{var} + \sum_{i=1}^{n_{var}} (u_i^2 - 10 \cos(2\pi u_i))$	$n_{var} = 20$; $-600 \leq u_i \leq 600$	0 at $u = (0, \dots, 0)$	Thousands of local minima
Griewank (GW $_{n_{var}}$), $F_{obj} = \sum_{i=1}^{n_{var}} u_i^2 / d - \prod_{i=1}^{n_{var}} \cos(u_i / \sqrt{i}) + 1$	$n_{var} = 20$; $-600 \leq u_i \leq 600$	0 at $u = (0, \dots, 0)$	Hundreds of local minima
Hartman (HAR $_{n_{var}}$), $F_{obj} = - \sum_{i=1}^4 c_i \exp \left[- \sum_{j=1}^{n_{var}} a_{ij} (u_j - p_{ij})^2 \right]$	$n_{var} = 3$ and 6 ; $0 \leq u_i \leq 1$	-3.862782 at $u = (0.114614, 0.555649, 0.852547)$ for $n_{var} = 3$; -3.322368 at $u = (0.201690, 0.150011, 0.476874, 0.275332, 0.311652, 0.657301)$ for $n_{var} = 6$	4 local minima, parameters c_i , a_{ij} and p_{ij} are reported by Ali et al. [19]
Shekel (SHE $_m$), $F_{obj} = - \sum_{i=1}^m \frac{1}{\sum_{j=1}^4 (u_j - a_{ij})^2 + c_i}$ for $m = 5, 7$ and 10	$n_{var} = 4$; $0 \leq u_i \leq 10$	-10.15 at $u = (4, \dots, 4)$ for $m = 5$; -10.40 at $u = (4, \dots, 4)$ for $m = 7$; -10.53 at $u = (4, \dots, 4)$ for $m = 10$	m local minima, parameters c_i and a_{ij} are available in Ali et al. [19]

Table 2
Suggested values of parameters in the stochastic optimization methods for solving benchmark and VLE data modeling problems.

Method	Parameter	Tested values	Suggested value	
			Benchmark problems	VLE data modeling problems
SA	$T_{SA,0}$	0.1–1000	10	10
	$T_{SA,F}$	10^{-9} – 10^{-3}	10^{-6}	10^{-6}
GA	P_{cros}	0.5–0.9	0.8	0.8
	P_{mut}	0.001–0.5	0.5	0.1
PSO	$C_{1,0}$	1.5–3.5	3.5	3.5
	w_0	0.7–0.9	0.8	0.8
DE	CR	0.1–0.9	0.5	0.1
	A	0.1–0.7	0.5	0.1
DETL	CR	0.1–0.7	0.5	0.3
	A	0.1–0.9	0.5	0.1
	tr	$0.0001n_{var}$ – $0.01n_{var}$	$0.001n_{var}$	$0.001n_{var}$
	tls	5–20	20	20

tially non-convex with several local minima within the specified bounds. Previous studies have shown that parameter estimation in local composition models for VLE data modeling involves solution of a global optimization problem [1,3,4,9–11]. In this study we have tested SA, GA, DE, DETL and PSO for the global minimization of the objective functions (Eqs. (9) or (15)) for VLE data modeling.

4. Benchmark problems

In addition to the VLE data modeling problems, several benchmark problems having 2–20 decision variables and different number of local optima, were used to evaluate and compare the five stochastic methods, each followed by the quasi-Newton method. Performance on benchmark problems is used as the starting point to test the optimization strategies, to gain insights on how the algorithms behave and to analyze their relative merits for well-known objective functions. Further, the results on benchmark problems help to generalize the relative performance of the optimization techniques. We have selected some benchmark problems considering the characteristics and difficulty of VLE data modeling problems (e.g., multimodal and multivariable problems such as those of error-in-variable formulation). So, the results are useful to identify the potential strengths and weaknesses of the tested stochastic optimization methods for VLE data modeling.

Table 1 provides the details of the benchmark functions, which have been used for testing stochastic optimization methods in other studies, e.g. [16,19,30,34]. The performance of all stochastic methods is evaluated based on both reliability (measured in terms of number of times the algorithm located the global minimum out of 100 trials, referred as success rate SR) and computational efficiency (measured in terms of average number of function evaluations NFE and CPU time). Note that NFE includes both the function call for evaluating the objective function using the stochastic method (NFE_{stc}) and the function calls for the local optimization (NFE_{qn}). The average NFE and CPU time are evaluated using successful trials only. 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.

4.1. Parameter tuning

Rosenbrock, Griewank and Shekel functions (Table 1) have been used to establish the most suitable parameter values for solving the benchmark problems efficiently and reliably. Parameter tuning was carried out by varying one parameter at a time with the remain-

ing parameters fixed at nominal values, which were established using values reported in the literature [16,23,24,28,30]. The tested and suggested values for parameters of each stochastic method are summarized in Table 2.

4.2. Results and discussion

All stochastic methods were studied using the two stopping criteria: (a) maximum number of iterations/generations $Iter_{max}$ or Gen_{max} (referred as stopping criterion 1, SC1) and (b) maximum number of iterations/generations without improvement in the best function value Sc_{max} (referred as stopping criterion 2, SC2). The performance of all stochastic methods is compared in terms of SR and NFE by examining different levels of algorithm efficiency, which are obtained by changing the values of $NS \times NT$ ($=NP = n_p = N_{pop}$), $Iter_{max}/Gen_{max}$ and Sc_{max} . Note that optimal values of these parameters may be problem dependent, and also determine the trade-off between efficiency and reliability. As a consequence, selection of proper values for them is important for the comparison.

The results obtained for benchmark problems are presented in Fig. 1 and Tables 3–6. For an overall comparison of algorithms, we report the global success rate (GSR) of a stochastic optimization method, which is defined as the mean success rate for all benchmark problems tested (nb), i.e. $GSR = \sum_{i=1}^{nb} (SR_i/nb)$ where SR_i is the success rate in the problem i . The results indicate that the GSR

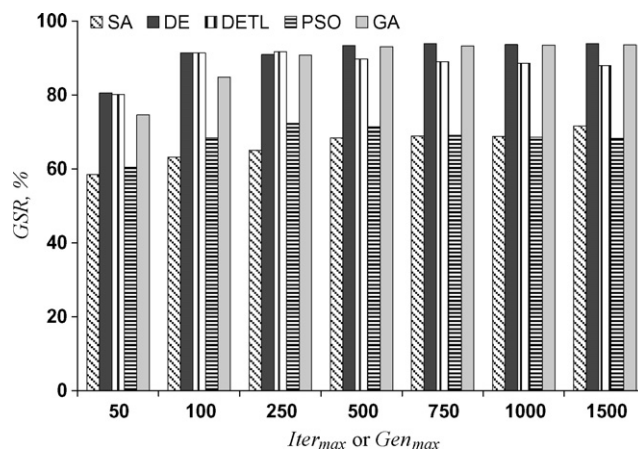


Fig. 1. Global success rate (GSR) versus $Iter_{max}/Gen_{max}$ (without using Sc_{max}) of SA, DE, DETL, PSO and GA for benchmark problems. Algorithm parameters: $NS \times NT = NP = n_p = N_{pop} = 10n_{var}$.

Table 3
Global success rate (GSR) versus Sc_{max} of SA, DE, DETL, PSO and GA for benchmark and VLE modeling problems. Algorithm parameters: $NS \times NT = NP = n_p = N_{pop} = 10n_{var}$.

Method	Sc_{max}^a	GSR (%) for		
		Benchmark	VLE-LS	VLE-EIV
SA	$6n_{var}$	64	14	35
	$12n_{var}$	65	37	38
DE	$6n_{var}$	92	3	67
	$12n_{var}$	94	12	67
DETL	$6n_{var}$	89	15	63
	$12n_{var}$	90	33	61
PSO	$6n_{var}$	65	2	0
	$12n_{var}$	65	7	4
GA	$6n_{var}$	59	4	0
	$12n_{var}$	70	9	0

^a $Iter_{max}/Gen_{max}$ is restricted to a maximum of 1500.

of DE and DETL is superior to other methods tested for solving the benchmark problems using either SC1 or SC2 as the stopping condition (see Fig. 1 and Table 3). GSR of DE and DETL is higher than 80% and the number of successes increased as the $Iter_{max}/Gen_{max}$ and Sc_{max} increased. Note that GA outperformed the SA and PSO in solving benchmark problems especially for SC1, with a GSR > 75%. It appears that SA and PSO provide very similar GSR, and they are the worse performers especially if SC1 is used alone as the stopping condition. Though SA and PSO showed the worst performance, their success rate can be still considered competitive for several benchmark problems. In particular, both SA and PSO showed a high success rate for Zakharov, HAR₃, ROS₁₀, ROS₅ and ROS₂ functions.

Table 4
NFE of SA/GA/PSO/DE for solving benchmark problems with SC1 alone as the stopping criterion.

F_{obj}	n_{var}	NFE for $Iter_{max}$ or Gen_{max}^a						
		50	100	250	500	750	1000	1500
ROS ₂	2	1,039	2,035	5,028	10,026	15,026	20,025	30,025
HAR ₃	3	1,556	3,055	7,551	15,051	22,550	30,051	45,051
SHE ₅	4	2,083	4,081	10,069	20,064	30,063	40,063	60,062
ROS ₅	5	2,668	5,156	12,637	25,123	37,613	50,108	75,102
HAR ₆	6	3,196	6,189	15,167	30,171	45,168	60,168	90,167
ROS ₁₀	10	5,414	10,479	25,438	50,375	75,357	100,341	150,321
ROS ₂₀	20	10,934	21,283	51,343	101,243	151,255	201,093	301,048

^a Algorithm parameters: $NS \times NT = NP = n_p = N_{pop} = 10n_{var}$.

Table 5
Percent reduction in NFE of DETL for solving benchmark problems with SC1 alone as the stopping criterion.

F_{obj}	Percent reduction in NFE for $Iter_{max}$ or Gen_{max}^a						
	50	100	250	500	750	1000	1500
ZAK ₂₀	1.48	0.64	0.33	0.16	0.10	0.08	0.05
ZAK ₁₀	1.65	0.88	0.30	-0.78	-20.82	-34.17	-48.88
ZAK ₅	1.57	0.70	-37.73	-60.69	-73.88	-78.62	-88.96
ZAK ₂	-34.79	-59.26	-76.76	-83.83	-86.16	-87.33	-88.96
ROS ₂₀	-	2.59	0.65	-2.00	-5.89	-9.44	-14.31
ROS ₁₀	2.62	1.03	0.14	-0.93	-2.27	-3.68	-4.09
ROS ₅	1.14	0.25	0.13	-0.31	-0.83	-4.78	-17.65
ROS ₂	0.57	-3.90	-41.64	-64.45	-72.73	-77.54	-82.82
GP	-16.87	-50.21	-72.82	-81.43	-84.98	-86.37	-88.74
mHB	-7.56	-42.18	-69.50	-79.59	-83.91	-85.97	-88.44
RAS ₂₀	-	-	-	-	-	-	-
GW ₂₀	1.49	0.66	-14.25	-41.53	-52.67	-58.41	-64.84
HAR ₃	-14.49	-48.51	-71.56	-79.53	-83.17	-84.40	-87.00
HAR ₆	0.95	-8.09	-42.66	-58.73	-64.62	-62.49	-65.02
SHE ₅	1.92	-1.15	-24.26	-46.28	-59.56	-61.72	-67.84
SHE ₇	1.78	-3.40	-31.76	-54.03	-63.46	-67.83	-70.99
SHE ₁₀	1.81	-2.33	-34.86	-55.80	-64.94	-68.52	-74.06

^a % reduction = 100(NFE of DETL - mean NFE of other stochastic methods tested)/mean NFE of other stochastic methods tested. The symbol “-” indicates that NFE is not reported because the stochastic method showed a 0% SR.

By examining performance at different levels of computational effort, our data indicate that all methods may fail to locate the global optimum depending on the algorithm parameters; see results reported in Fig. 1 and Table 3. These failures are due to the presence of local optima in most of the benchmark problems and due to the narrow global minimum region of some challenging objective functions [16]. In some multivariable problems, the SR of all methods decreases due to the increase in the complexity of solution space. As indicated by Ali et al. [19], problems with higher dimension are expected to be more difficult to solve than those with lower dimension. In general, the reliability and computational effort of the methods increase with $NS \times NT (=NP = n_p = N_{pop})$, $Iter_{max}/Gen_{max}$ and Sc_{max} . However, the reliability of the methods may be poor when the $Iter_{max}/Gen_{max}$ is limited to low values especially for functions with many decision variables.

Comparing the performance of stochastic methods, each followed by the quasi-Newton method, at the same level of computational effort using NFE as reference and SC1 as the stopping criterion, the GSR of DE and DETL is better than that obtained for SA, GA and PSO especially in early iterations (Fig. 1). But, as the number of generations/iterations increases, GA performs relatively well among the tested methods. Note that DE dominates the other algorithms over most of the iteration/generation values tested. DETL has also a good performance and may require lower NFE than the other stochastic methods using SC1 alone as the stopping criterion for the benchmark problems tested (see Tables 4 and 5). The NFE of DE, PSO, SA and GA is almost the same when SC1 is used as stopping criterion, as shown in Table 4. The percentage reduction in total NFE of DETL compared to other stochastic methods for SC1 is summarized in Table 5. Overall, a reduction from 0.3 to 89%

Table 6
NFE of SA, DE, DETL, PSO and GA for solving benchmark problems using SC2 alone as the stopping criterion.

F_{obj}	Sc_{max}	NFE for ^a				
		SA	DE	DETL	PSO	GA
ZAK ₂₀	6 n_{var}	300,235	103,138	129,553	57,265	73,060
	12 n_{var}	300,237	272,956	287,243	181,470	166,995
ZAK ₁₀	6 n_{var}	44,253	105,930	44,436	17,413	12,287
	12 n_{var}	124,614	150,117	54,955	82,311	33,499
ZAK ₅	6 n_{var}	5,597	38,214	6,701	4,400	2,867
	12 n_{var}	12,591	40,114	7,340	21,783	6,286
ZAK ₂	6 n_{var}	607	3,057	625	733	491
	12 n_{var}	1,159	5,064	681	1,747	989
ROS ₂₀	6 n_{var}	285,359	302,174	146,430	208,943	111,088
	12 n_{var}	301,206	302,157	237,659	299,870	271,409
ROS ₁₀	6 n_{var}	55,526	150,620	52,798	24,103	15,771
	12 n_{var}	131,195	150,619	115,299	107,527	38,203
ROS ₅	6 n_{var}	5,845	16,722	9,533	4,673	3,562
	12 n_{var}	14,495	50,768	22,394	12,876	6,425
ROS ₂	6 n_{var}	1,260	786	864	666	699
	12 n_{var}	1,991	3,345	1,636	1,122	1,066
GP	6 n_{var}	736	1,709	725	655	567
	12 n_{var}	1,218	3,637	907	1,409	970
mHB	6 n_{var}	683	1,796	731	613	527
	12 n_{var}	1,302	4,488	1,056	1,289	951
RAS ₂₀	6 n_{var}	–	–	–	–	–
	12 n_{var}	–	–	–	–	–
GW ₂₀	6 n_{var}	1,695	4,687	1,415	1,081	1,126
	12 n_{var}	2,921	5,293	1,566	2,038	2,114
HAR ₃	6 n_{var}	280,305	232,381	49,720	194,233	141,097
	12 n_{var}	300,842	257,391	59,817	301,507	297,881
HAR ₆	6 n_{var}	2,511	6,891	4,165	1,825	1,944
	12 n_{var}	5,510	12,674	7,237	3,620	3,674
SHE ₅	6 n_{var}	2,431	7,943	4,410	1,860	1,931
	12 n_{var}	5,382	12,744	5,672	3,469	3,815
SHE ₇	6 n_{var}	2,627	8,282	4,267	1,768	2,037
	12 n_{var}	5,616	12,177	5,901	3,536	3,532
SHE ₁₀	6 n_{var}	6,291	19,163	7,041	3,981	4,096
	12 n_{var}	13,864	21,782	11,105	7,840	8,624

^a $Iter_{max}/Gen_{max}$ is restricted to a maximum of 1500. Algorithm parameters: $NS \times NT = NP = n_p = N_{pop} = 10n_{var}$. The symbol “–” indicates that NFE is not reported because the stochastic method showed a 0% SR.

in NFE of DETL is achieved in the benchmark problems tested. As stated earlier, DETL is a hybrid strategy that integrates the strong features of DE and TS. The tabu concept has considerable influence on the performance of DETL resulting in lower NFE [16]. However, the improvement in the convergence rate of DETL may not be the same for all the functions and depends on problem dimensionality and complexity.

If SC2 is used alone as the convergence criterion (along with $Iter_{max}/Gen_{max}$ set at 1500 for keeping the computational effort within reasonable values), the results indicate that DE and DETL can achieve a higher SR than those obtained for GA, PSO and SA (see Table 3). At tested values of SC2, SA, PSO and GA gave similar GSR for benchmark problems. It is interesting that the NFE of GA is lower than those of SA, DE, DETL and PSO using SC2 especially for multi-variable problems as shown in Table 6. For this stopping condition, DE generally required more NFE than the other algorithms.

When we examine the results for individual problems, all stochastic methods show high reliability for Zakharov and HAR₃ functions irrespective of the stopping criterion used (i.e., $Iter_{max}/Gen_{max}$ or Sc_{max}). In fact, the global optimum is found even using low values of these stopping conditions. DE and DETL achieved high SR values and their performance is usually better

than that of GA, SA and PSO for both SC1 and SC2 in almost all benchmark problems. If SC1 is used as the stopping condition, the reliability of SA is low compared to DE, DETL, GA and PSO for functions ROS₁₀, ROS₅, GP₂, mHB and GW₂₀, and PSO has the lowest SR for functions ROS₂₀, SHE₅, SHE₇, SHE₁₀ and HAR₆. In the case of RAS₂₀, the tested stochastic methods performed poorly; this indicates that they are affected by the huge number of local minima present in this challenging benchmark problem. For difficult benchmark problems, the performance of all stochastic methods could be improved using larger $Iter_{max}/Gen_{max}$ (i.e. >1500) or Sc_{max} (i.e. >12 n_{var}) but at the expense of a significant computational effort. With respect to the stopping condition SC2, the SR of SA, on average, is low for GP, mHB and GW₂₀ whereas PSO is the worst performer for ROS₂₀, SHE₅, SHE₇, SHE₁₀ and HAR₆. Further, GA failed to find the global optimum in ROS₂, ROS₅ and ROS₁₀ several times.

In the tests performed on the benchmark problems, the CPU time ranged from 0.02 to 0.4 s for SA, from 0.04 to 3.3 s for DE, from 0.07 to 3.9 s for DETL, from 0.03 to 3.1 s for PSO, and from 0.04 to 3.0 s for GA, respectively. Obviously, it increases as NFE increases for all stochastic methods. SA took less CPU time compared to the population-based DE, DETL, PSO and GA because one iteration of SA implies less memory requirements and arithmetic operations com-

pared to the requirements of one generation for population-based methods. Though DETL may require less NFE especially for SC1, it took more CPU time compared to the other four stochastic methods studied. As noted by Srinivas and Rangaiah [16], the diversification stage associated with tabu concept in DETL requires additional computational effort especially for problems such as benchmark functions where each function evaluation requires insignificant computational time. In summary, our results indicate that DE and DETL offer the best compromise between reliability and efficiency for solving the benchmark problems used in this study.

5. Parameter estimation problems in VLE data modeling

In this study, the five stochastic methods are tested and compared for several sets of binary VLE data, which have been studied by Gau et al. [3], Alvarez et al. [4], Gau and Stadtherr [6,7], Dominguez et al. [9], Srinivas and Rangaiah [10], Bonilla-Petriciolet et al. [11] and Bollas et al. [13]. These authors have used both LS and EIV formulations, local composition models and deterministic or stochastic global optimization methods for solving parameter estimation problems in VLE data modeling. Details of all examples (i.e., conditions of experimental data, thermodynamic models, objective function, decision variables and global optimum) are reported in Tables 7 and 8. All the experimental data are taken from DECHEMA collection. According to reported studies, there are at least two local minima (including the global minimum) in the specified interval for decision variables of all parameter estimation problems. The global optimum found for each example is in agreement with that reported in the literature. In short, the selected VLE problems are multimodal with dimension ranging from 2 to 20 decision variables. These problems have a variety of inherent difficulty and, as indicated, have been used to test different optimization strategies [3,4,6,7,9–11,13]. Therefore, the number and features of the selected VLE problems are sufficient to demonstrate and compare the performance of the tested stochastic methods.

5.1. Parameter tuning

Two examples were used for parameter tuning of all stochastic optimization methods. First example is the VLE data for the binary system tert butanol+1 butanol at 100 mmHg. This system was studied by Gau et al. [3] using interval analysis and classical least squares formulation, while Alvarez et al. [4] and Bonilla-Petriciolet et al. [11] have applied GA and SA for solving this parameter estimation problem. We optimize Eq. (9) with respect to the Wilson model parameters inside the interval: θ_1 and $\theta_2 \in (-8500, 320000)$. The initial values for each calculation and for all stochastic methods are randomly generated within these bounds. The vapor pressure is calculated by Antoine equation using the model parameters reported in Gau et al. [3].

The second example refers to modeling of VLE data of the binary system benzene + hexafluorobenzene using the error-in-variable formulation. This system has been studied, for example, by Gau and Stadtherr [7] using interval analysis whereas Srinivas and Rangaiah [10] and Bonilla-Petriciolet et al. [11] used the RTA and SA, respectively. One data set at 500 mmHg, which includes experimental data from 16 experiments, is used, and the data modeling is performed using the Wilson equation for liquid-phase activity coefficients and ideal gas model. Following Gau and Stadtherr [7], a standard deviation of (0.003, 0.0029, 1.7, 0.083) is assumed for state variables (x_1, y_1, P, T). The objective function is defined by Eq. (15) and it is optimized with respect to 34 decision variables. The initial intervals on the independent state variables $\bar{z} = (\bar{x}_{ij}, \bar{T})$ are set using plus and minus three standard deviations ($\pm 3\sigma$) while the intervals for the Wilson model parameters are defined as $\theta_1, \theta_2 \in (-10000, 200000)$.

Table 7
Details of VLE data modeling problems used for performance comparison of stochastic optimization methods, in the classical least squares formulation.

No.	System	Data	Models	Decision variables	Global optimum	Remarks
1	tert Butanol + 1 butanol	$P = 100$ mmHg, $ndat = 9$	Wilson and Ideal gas model	$n_{var} = 2; \theta_1, \theta_2 \in (-8500, 320000)$	$F_{obj} = 0.01026, \theta_1 = -567.96, \theta_2 = 745.33$	2 minima: Gau et al. [3], Alvarez et al. [4], Bonilla-Petriciolet et al. [11]
2		$P = 700$ mmHg, $ndat = 9$			$F_{obj} = 0.013690, \theta_1 = -733.95, \theta_2 = 1318.23$	
3		$P = 500$ mmHg, $ndat = 9$			$F_{obj} = 0.006852, \theta_1 = -718.01, \theta_2 = 1264.74$	
4	Water + 1,2 ethanediol	$P = 430$ mmHg, $ndat = 18$	Wilson and Ideal gas model	$n_{var} = 2; \theta_1, \theta_2 \in (-8500, 320000)$	$F_{obj} = 1.039134, \theta_1 = 5072.361, \theta_2 = -1921.62$	
5			UNIQUAC and Ideal gas model	$n_{var} = 2; \theta_1, \theta_2 \in (-5000, 20000)$	$F_{obj} = 1.408547, \theta_1 = -1131.84, \theta_2 = 3617.65$	
6			NRTL and Ideal gas model	$n_{var} = 3; \theta_1, \theta_2 \in (-2000, 5000); \alpha_{12} \in (0.01, 10.0)$	$F_{obj} = 1.253531, \theta_1 = -678.99, \theta_2 = 3046.13, \alpha_{12} = 0.621375$	
7	Benzene + hexafluorobenzene	$T = 50$ °C, $ndat = 11$	Wilson and Ideal gas model	$n_{var} = 2; \theta_1, \theta_2 \in (-8500, 320000)$	$F_{obj} = 0.008935, \theta_1 = -424.08, \theta_2 = 983.06$	2 minima: Gau et al. [3], Bonilla-Petriciolet et al. [11], Bollas et al. [13]
8		$P = 300$ mmHg, $ndat = 17$			$F_{obj} = 0.014860, \theta_1 = -432.49, \theta_2 = 992.85$	
9		$P = 760$ mmHg, $ndat = 29$			$F_{obj} = 0.014616, \theta_1 = -334.70, \theta_2 = 704.74$	
10		$T = 30$ °C, $ndat = 10$			$F_{obj} = 0.011783, \theta_1 = -467.76, \theta_2 = 1313.94$	

Table 8 Details of VLE data modeling problems used for performance comparison of stochastic optimization methods, in the error-in-variable formulation.

No.	System	Data	Models	Decision variables	Global optimum	Remarks
11	Benzene + hexafluorobenzene	$P = 500 \text{ mmHg}$, $n_{\text{dat}} = 16$, $\sigma(0.003, 0.0029, 1.7, 0.083)$	Wilson and Ideal gas model	$n_{\text{var}} = 34$; $\theta_1, \theta_2 \in (-10000, 200000)$; $x_{ij}^c \in (x_{ij} - 3\sigma, x_{ij} + 3\sigma)$; $T^c \in (T - 3\sigma, T + 3\sigma)$ $n_{\text{var}} = 36$	$F_{\text{obj}} = 19.998720$, $\theta_1 = -429.85$, $\theta_2 = 1029.32$	2 minima; Gau and Stadtherr [6,7], Srinivas and Rangaiah [10], Bonilla-Petriciolet et al. [11]
12		$P = 300 \text{ mmHg}$, $n_{\text{dat}} = 17$, $\sigma(0.003, 0.0029, 1.7, 0.083)$		$n_{\text{var}} = 22$	$F_{\text{obj}} = 42.343724$, $\theta_1 = -437.72$, $\theta_2 = 1003.12$	
13		$T = 30^\circ\text{C}$, $n_{\text{dat}} = 10$, $\sigma(0.001, 0.01, 0.75, 0.1)$		$n_{\text{var}} = 22$	$F_{\text{obj}} = 11.898795$, $\theta_1 = -472.00$, $\theta_2 = 1274.32$	
14		$T = 40^\circ\text{C}$, $n_{\text{dat}} = 10$, $\sigma(0.001, 0.01, 0.75, 0.1)$		$n_{\text{var}} = 22$	$F_{\text{obj}} = 11.170496$, $\theta_1 = -462.51$, $\theta_2 = 1197.31$	
15		$T = 50^\circ\text{C}$, $n_{\text{dat}} = 11$, $\sigma(0.003, 0.0029, 1.7, 0.083)$		$n_{\text{var}} = 24$	$F_{\text{obj}} = 25.671042$, $\theta_1 = -415.76$, $\theta_2 = 944.57$	
16		$T = 60^\circ\text{C}$, $n_{\text{dat}} = 10$, $\sigma(0.003, 0.0029, 1.7, 0.083)$		$n_{\text{var}} = 22$	$F_{\text{obj}} = 19.401593$, $\theta_1 = -437.05$, $\theta_2 = 1065.04$	

As before, the parameters of each stochastic method are tuned one at a time while keeping others fixed at their nominal values. Based on results of parameter tuning, not given in this paper for brevity, the suggested values of parameters in the stochastic method for VLE data modeling are reported in Table 2. The suggested values of some parameters in GA, DE and DETL are slightly different from those suggested for benchmark problems due to the different characteristics of objective functions (e.g., only a few minima in the VLE parameter estimation problems and non-linearity).

5.2. 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 stochastic methods is compared using SR, NFE and CPU time for both stopping criteria—SC1 and SC2. The results of solving the VLE parameter estimation problems for different values of these stopping conditions are shown in Figs. 2–4 and Tables 3, 9, 10 and 11. In the first instance, to directly compare the performance of algorithms, we keep their numerical effort the same via $Iter_{\text{max}}/Gen_{\text{max}}$ (i.e., without using Sc_{max}) and compare the results obtained in terms of GSR. The performance of the methods tested is depicted graphically in Figs. 2 and 3 to illustrate their GSR as a function of $Iter_{\text{max}}/Gen_{\text{max}}$ for all VLE problems. These figures show that the reliability of the stochastic methods highly depends on and increases with $Iter_{\text{max}}/Gen_{\text{max}}$. As expected, the reliability of all the methods increases with $NS \times NT (=NP = n_p = N_{\text{pop}})$ as shown in Fig. 2b. The stochastic methods may fail in the global minimization of the objective functions especially for EIV problems. For illustration, plots of SR versus $Iter_{\text{max}}/Gen_{\text{max}}$ for selected VLE problems are given in Fig. 4.

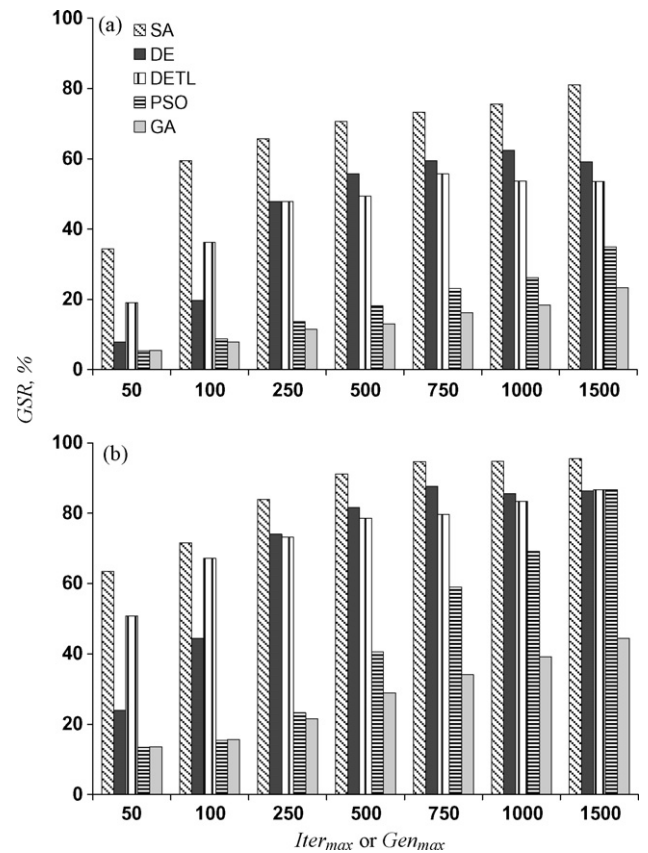


Fig. 2. Global success rate (GSR) versus $Iter_{\text{max}}/Gen_{\text{max}}$ (without using Sc_{max}) of SA, DE, DETL, PSO and GA for VLE data modeling problems using LS formulation: (a) $NS \times NT = NP = n_p = N_{\text{pop}} = 10n_{\text{var}}$; and (b) $NS \times NT = NP = n_p = N_{\text{pop}} = 50n_{\text{var}}$.

Table 9
NFE of SA, DE, PSO and GA for solving VLE modeling problems using SC1 alone as the stopping criterion; algorithm parameters: $NS \times NT = NP = n_p = N_{pop} = 10n_{var}$.

No.	n_{var}	NFE for $Iter_{max}$ or Gen_{max}						
		50	100	250	500	750	1000	1500
1	2	1,071	2,063	5,067	10,063	15,064	20,066	30,065
6	3	1,685	3,238	7,729	15,222	22,712	30,211	45,209
11	34	18,661	36,201	92,953	179,242	264,725	354,938	524,500
12	36	19,817	38,492	97,994	193,175	283,661	374,183	556,462
13	22	11,991	23,016	56,515	113,674	169,235	224,882	338,271
15	24	13,110	25,086	62,899	129,606	192,706	253,423	371,940

Table 10
Percent reduction in the NFE of DETL for solving VLE modeling problems using SC1 alone as the stopping criterion.

No.	Percent reduction in NFE for $Iter_{max}$ or Gen_{max} ^a						
	50	100	250	500	750	1000	1500
1	-11.84	-24.21	-51.17	-63.11	-68.36	-70.29	-74.28
2	-	-38.52	-59.35	-71.62	-75.75	-78.57	-80.85
3	-35.36	-41.26	-60.20	-71.40	-74.46	-78.26	-82.08
4	-7.79	-27.19	-41.52	-54.68	-55.58	-58.46	-63.71
5	-6.77	-30.62	-56.05	-68.97	-73.82	-77.34	-78.33
6	0.77	-1.69	-14.93	-39.11	-47.15	-58.94	-66.06
7	-12.45	-20.66	-40.80	-55.29	-62.09	-67.35	-68.68
8	-9.28	-25.24	-51.94	-63.84	-71.61	-74.20	-75.75
9	-7.78	-16.42	-39.80	-52.88	-58.89	-63.11	-67.02
10	-10.96	-22.83	-50.67	-65.97	-68.96	-71.56	-75.45
11	-	-	3.97	-10.60	-30.07	-39.62	-42.79
12	-	-	-	-	-	-	-
13	-	-	-	-	-	-	-
14	-	-	5.48	0.03	-0.50	-8.18	-24.63
15	-	-	-	-11.30	-32.43	-42.77	-49.93
16	-	-	-	-10.01	-30.26	-40.22	-45.90

^a % reduction = $100(\text{NFE of DETL} - \text{mean NFE of other stochastic methods tested}) / \text{mean NFE of other stochastic methods tested}$. The symbol “-” indicates that NFE is not reported because the stochastic method showed a 0% SR.

The reliability of SA is better than that of DE, DETL, PSO and GA for tested VLE examples using LS formulation and SC1. It is interesting to observe that the behavior of each stochastic method in VLE problems using LS formulation is consistent. Specifically, SA generally outperformed other stochastic methods for solving VLE problems in terms of success rate throughout the tested range of $Iter_{max}/Gen_{max}$. Only for VLE problems No. 6 and 9, SA showed 100% reliability for the global minimization of the objective function if proper values of $Iter_{max}$ are used. The SR of SA ranged from 9 to 100% in VLE problems No. 1–10. On the other hand, DE and DETL can perform reasonably well in VLE problems using LS formulation (GSR \cong 86%), but they showed the best performance for problems based on EIV formulation and using SC1 alone as the stopping con-

dition. Specifically, Fig. 3 shows that the GSR of SA is better than those obtained for DE and DETL in early iterations for EIV problems, and vice versa as the number of generations/iterations increases. These results suggest that SA may be preferred if a small number of function evaluations (i.e. fewer iterations/generations) is allowed to solve EIV problems using the SC1 as the convergence criterion. Otherwise, DE and DETL are more suitable if a higher number of function evaluations are permitted for this stopping condition. The exceptions are the EIV problems No. 13 and 14 for which the SR of SA is less than 20% but better than that obtained for DE (SR \leq 6%) and DETL (SR \leq 1%) in the range of iterations/generations tested. In particular, these VLE problems involve challenging global optimization functions and illustrate the difficulty of locating the globally optimal parameters for VLE data modeling by the EIV approach. The maximum GSRs of DE and DETL are 66 and 63% for VLE problems No. 11–16, whereas SA showed a maximum GSR of 39% in these problems (Fig. 3).

Surprisingly, GA performed worse than all other stochastic methods tested for both LS and EIV formulations using SC1 as the stopping condition. GA may achieve a maximum GSR of 45% for LS problems and 1% for EIV problems at tested conditions. It appears that GA is frequently trapped by the local minima of the objective functions of selected VLE problems especially for EIV formulation. 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 [3]. Upon applying PSO with SC1, the performance results were better than those of GA in terms of GSR for both LS and EIV formulations (Figs. 2 and 3). The maximum GSR obtained for PSO was 87% for LS problems and 10% for EIV problems, respectively. Results reported in Fig. 2b indicate that PSO is competitive if a higher value of N_{pop} (e.g., $50n_{var}$) is used and its reliability may be comparable to that of DE and DETL. But, on average, PSO is less

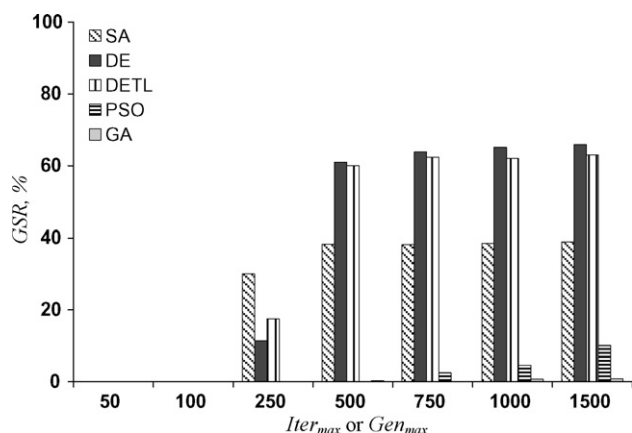


Fig. 3. Global success rate (GSR) versus $Iter_{max}/Gen_{max}$ (without using Sc_{max}) of SA, DE, DETL, PSO and GA for VLE data modeling problems using EIV formulation. Algorithm parameters: $NS \times NT = NP = n_p = N_{pop} = 10n_{var}$.

Table 11
NFE of stochastic optimization methods for solving VLE modeling problems using SC_{max} alone as the stopping condition (SC2).

No.	SC_{max}	NFE for ^a				
		SA	DE	DETL	PSO	GA
1	$6n_{var}$	786	–	877	–	–
	$12n_{var}$	1,892	1,311	1,777	–	1,063
2	$6n_{var}$	734	–	902	–	585
	$12n_{var}$	1,607	1,664	1,718	–	990
3	$6n_{var}$	762	–	882	–	–
	$12n_{var}$	1,750	1,557	1,592	–	1,054
4	$6n_{var}$	715	535	928	–	568
	$12n_{var}$	1,087	1,711	1,564	1,047	1,043
5	$6n_{var}$	862	632	830	611	669
	$12n_{var}$	1,497	1,678	1,327	1,205	1,036
6	$6n_{var}$	2,269	1,923	3,024	1,623	1,841
	$12n_{var}$	3,625	4,003	5,218	2,520	2,531
7	$6n_{var}$	704	557	927	–	582
	$12n_{var}$	1,317	1,353	1,900	–	1,064
8	$6n_{var}$	871	–	1,002	–	592
	$12n_{var}$	1,578	1,583	1,756	–	1,101
9	$6n_{var}$	730	–	854	–	–
	$12n_{var}$	1,617	1,900	1,911	1,021	1,036
10	$6n_{var}$	836	–	1,015	–	579
	$12n_{var}$	1,664	1,561	1,653	–	1,005
11	$6n_{var}$	530,150	520,573	299,739	–	–
	$12n_{var}$	530,803	520,549	298,581	397,855	–
12	$6n_{var}$	323,665	350,029	–	–	–
	$12n_{var}$	330,439	330,855	–	–	–
13	$6n_{var}$	330,473	339,809	–	–	–
	$12n_{var}$	330,487	374,026	–	–	–
14	$6n_{var}$	552,509	552,966	274,999	–	–
	$12n_{var}$	551,682	553,199	279,205	453,666	–
15	$6n_{var}$	368,086	372,382	178,550	–	–
	$12n_{var}$	367,988	371,841	180,862	244,653	143,024
16	$6n_{var}$	331,555	335,159	171,136	–	–
	$12n_{var}$	341,566	339,205	182,271	186,245	–

^a $Iter_{max}$ and Gen_{max} are restricted to a maximum value of 1500. Algorithm parameters: $NS \times NT = NP = n_p = N_{pop} = 10n_{var}$. The symbol “–” indicates that NFE is not reported because the stochastic method showed a 0% SR.

successful than DE and DETL for both types of parameter estimation problems. Overall, GA and PSO performed poorly in VLE data modeling problems with EIV formulation.

The other three stochastic methods tested may reach a high reliability (i.e., SR > 70%) if proper values of algorithm parameters are used, in VLE examples using SC1 and both LS and EIV formulations. However, this good performance may be at the expense of computational effort for SA and DE, whereas DETL can achieve a high SR using lower NFE (Tables 9 and 10). The computational efficiency of DETL is usually better than that of SA, GA, PSO and DE for both LS and EIV formulations and SC1. Specifically, the percentage reduction in NFE of DETL ranged from 2 to 82% for LS functions and from 0.5 to 50% for EIV problems compared to other stochastic methods tested.

The performance of five stochastic methods for VLE examples No. 1–16 but using SC2 alone as a stopping condition is reported in Tables 3 and 11. As before, in these tests, $Iter_{max}/Gen_{max}$ is set at 1500. The results indicate that the GSR of SA and DETL is better compared to DE, PSO and GA for this convergence criterion and using LS formulation, whereas DE and DETL offer the best reliability for EIV problems. Note that, even though EIV problem involves many decision variables, DE and DETL showed a high SR for problems No. 11 and 14–16 using either SC1 or SC2 as the stopping condition. The performance of stochastic methods improves as SC2 increases in all

VLE problems. However, all methods showed poor performance in VLE problems No. 13 and 14 (i.e., the best SR was 16% for SA). With respect to NFE, it appears that the convergence rate of GA is faster compared to SA, DE, DETL and PSO using SC2. Unfortunately, GA showed generally low SR for solving VLE problems using both LS and EIV formulations.

Our numerical experience indicates that the reliability of all the stochastic methods is better using SC1 compared to that of SC2. This may be because of the slow convergence of some stochastic methods, which may require more iterations/generations to escape from the local minimum region reached in the initial iterations. However, the maximum number of generations/iterations to find the global optimum cannot be judged *a priori* for an arbitrary function. In some problems, this may lead to unnecessary function calls when the minimum is reached long before the maximum number of generations/iterations, thus increasing computational effort. Therefore, to avoid the premature convergence and favor the performance of stochastic methods, it would be useful to simultaneously employ a combination of suitable values for both SC1 and SC2 as stopping conditions (i.e., the algorithm terminates after satisfying either $Iter_{max}/Gen_{max}$ or SC_{max}).

The CPU time needed to perform the global optimization for the VLE problems considered ranged from 0.02 to 13.5 s for SA, from 0.47 to 23.7 s for PSO, from 0.02 to 22.9 s for DE, from 0.02 to 18.9 s

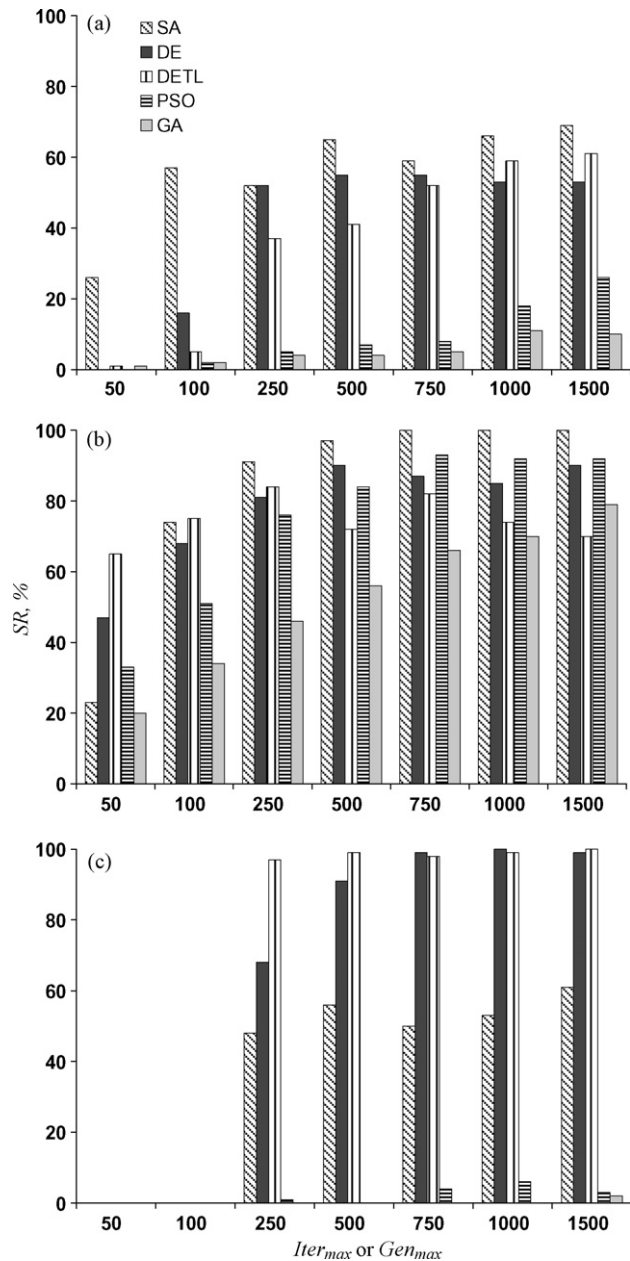


Fig. 4. Success rate (SR) versus $Iter_{max}/Gen_{max}$ (without using Sc_{max}) of SA, DE, DETL, PSO and GA for selected VLE problems using LS and EIV formulation: (a) No. 3, (b) No. 6 and (c) No. 11. Algorithm parameters: $NS \times NT = NP = n_p = N_{pop} = 10n_{var}$.

for DETL, and from 0.02 to 22.5 s for GA, respectively. The computing time of stochastic methods increases proportional to problem dimensionality (i.e., number of decision variables). However, it appears that for large NFE and multivariable problems similar to those tested, both SA and DETL are more efficient in terms of CPU time than other stochastic methods.

Finally, our study shows that the reliability and efficiency of the different stochastic optimization methods vary significantly for different VLE modeling problems and depends on the stopping criterion used. In general, VLE modeling using both LS and EIV formulations is a difficult global optimization problem and their characteristics pose a challenge to any optimization technique. For the many tests performed, DE and DETL are among the best algorithms and are suitable methods for parameter estimation in VLE modeling using both LS and EIV formulations. Note that this conclusion is consistent with the results reported for benchmark

problems, where both DE and DETL offer the best performance for global optimization of tested functions.

6. Conclusions

In this study, the performance of SA, GA, DE, DETL and PSO has been tested and compared for VLE modeling using experimental data for binary systems and both least squares and maximum likelihood criterions. Prior to this application, the five methods have been tested on benchmark problems with multiple minima. The performance of the stochastic optimization methods tested varies significantly between different problems and is dependent on the stopping criterion used, problem dimensionality and difficulty. For benchmark problems, DETL and DE are better than the other methods tested in terms of reliability; on the other hand, SA required the least CPU time while DETL required the least NFE. For VLE parameter estimation using the least squares criterion, GA is the worst performer whereas SA, DE, DETL and PSO show acceptable performance. Owing to the increase of problem dimensionality and difficulty in the EIV approach, all stochastic methods face difficulties in finding the global minimum for VLE data modeling. In some EIV problems, the performance of all stochastic methods tested is poor and not satisfactory. Also, there is a significant increase in the computational effort for the data fitting, caused by problem dimensionality. Overall, DE and DETL were found to perform better than other algorithms tested in terms of success rate for parameter estimation in VLE data modeling using both LS and EIV formulations. In particular, DETL offers a significant reduction in the computational time (both NFE and CPU time), which is attractive for solving VLE problems involving many parameters. In summary, results of this study show the strengths and weaknesses of several classical and recent stochastic global optimization methods for solving benchmark and VLE data modeling problems. They further show that DE and DETL are good alternatives and offer comparable or better performance than SA, PSO and GA methods for these global optimization problems. For further studies, we suggest the application and comparison of alternative stochastic optimization methods such as Ant Colony Optimization and Harmony Search for solving parameter estimation problems in phase equilibrium modeling.

List of symbols

A	amplification factor
C	number of components
C_1, C_2	cognitive and social parameters of Particle Swarm Optimization
CR	crossover constant
F_{obj}	objective function
Gen_{max}	maximum number of generations
$Iter_{max}$	maximum number of iterations
k	iteration counter
\vec{l}	continuous variable vector
NFE	number of function evaluations
N_{pop}	population size in Genetic Algorithm
NP	population size
NS	number of cycles of SA for updating decision variables
NT	number of iterations before annealing temperature reduction
n_p	swarm size (i.e., number of particles)
n_{var}	number of decision variables (i.e., parameters in parameter estimation problems)
n_h	neighborhood size in Particle Swarm Optimization
$ndat$	number of experiments
$npar$	number of parameters
$nest$	number of state variables
P	pressure

p_i^0	vapor pressure of pure component i
P_{cros}	crossover probability
P_{mut}	mutation probability
q_{ij}	Observation
r_{ij}	independent variables
R_i	random number
s_{ij}	position of particles in Particle Swarm Optimization
Sc_{max}	maximum number of successive iterations without improvement in the best function value
SR	success rate
tls	tabu list size
T	temperature
Tr	tabu radius
T_{SA}	annealing temperature of Simulated Annealing
$U_{i,G+1}$	trial vector
$V_{i,G+1}$	mutant vector
$V_{i,j}$	velocity of particles in Particle Swarm Optimization
V_{max}	maximum velocity of particles in Particle Swarm Optimization
VM	step length vector
W	inertia weight factor of Particle Swarm Optimization
x,y	mole fraction
$X_{i,G}$	target vector
z_{ij}^t	unknown “true” values of state variables
γ	activity coefficient
θ	parameter of thermodynamic model
σ_i	standard deviation associated with the measurement of state variable i

References

- [1] W.R. Esposito, C.A. Floudas, *Ind. Eng. Chem. Res.* 37 (1998) 1841–1858.
- [2] P. Englezos, N. Kalogerakis, *Applied Parameter Estimation for Chemical Engineers*, Marcel-Dekker, New York, 2001.
- [3] C.Y. Gau, J.F. Brennecke, M.A. Stadtherr, *Fluid Phase Equilib.* 168 (2000) 1–18.
- [4] V.H. Alvarez, R. Larico, Y. Ianos, M. Aznar, *Braz. J. Chem. Eng.* 25 (2008) 409–418.
- [5] E. Neau, A. Peneloux, *Fluid Phase Equilib.* 6 (1981) 1–19.
- [6] C.Y. Gau, M.A. Stadtherr, *AIChE J.* 48 (2002) 1192–1197.
- [7] C.Y. Gau, M.A. Stadtherr, in: R. Luus (Ed.), *Recent Developments in Optimization and Optimal Control in Chemical Engineering*, Transworld Research, 2002, pp. 1–17.
- [8] P. Englezos, N. Kalogerakis, P.R. Bishnoi, *Fluid Phase Equilib.* 61 (1990) 1–15.
- [9] A. Dominguez, J. Tojo, M. Castier, *Comp. Chem. Eng.* 26 (2002) 1473–1479.
- [10] M. Srinivas, G.P. Rangaiah, *Comp. Chem. Eng.* 30 (2006) 1400–1415.
- [11] A. Bonilla-Petriciolet, U.I. Bravo-Sanchez, F. Castillo-Borja, J.G. Zapaiain-Salinas, J.J. Soto-Bernal, *Braz. J. Chem. Eng.* 24 (2007) 151–162.
- [12] M. Kundu, A. Chitturi, S.S. Bandyopadhyay, *Can. J. Chem. Eng.* 86 (2008) 117–126.
- [13] G.M. Bollas, P.I. Barton, A. Mitsos, *Chem. Eng. Sci.* 64 (2009) 1768–1783.
- [14] A.L.H. Costa, F.P.T. da Silva, F.L.P. Pessoa, *Braz. J. Chem. Eng.* 17 (2000) 349–353.
- [15] F. Steyer, K. Sundmacher, *J. Chem. Eng. Data* 49 (2004) 1675–1681.
- [16] M. Srinivas, G.P. Rangaiah, *Ind. Eng. Chem. Res.* 46 (2007) 3410–3421.
- [17] I. Rahamn, A.Kr. Das, R.B. Mankar, B.D. Kulkarni, *Fluid Phase Equilib.* 282 (2009) 65–67.
- [18] Y.M. Cheng, L. Li, S.C. Chi, *Comput. Geotech.* 34 (2007) 462–484.
- [19] M.M. Ali, C. Khompatraporn, Z.B. Zabinsky, *J. Global Optim.* 31 (2005) 635–672.
- [20] S. Kirkpatrick, C. Gelatt, M. Vecchi, *Science* 220 (1983) 671–680.
- [21] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, *J. Chem. Phys.* 21 (1953) 1087–1092.
- [22] A. Corana, M. Marchesi, C. Martini, S. Ridella, *ACM T. Math. Software* 13 (1987) 262–280.
- [23] G.P. Rangaiah, *Fluid Phase Equilib.* 83 (2001) 187–188.
- [24] A. Bonilla-Petriciolet, R. Vázquez-Román, G.A. Iglesias-Silva, K.R. Hall, *Ind. Eng. Chem. Res.* 45 (2006) 4764–4772.
- [25] B. Goffe, G. Ferrier, J. Rogers, *J. Econometrics* 60 (1994) 65–99.
- [26] J.H. Holland, *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, 1975.
- [27] J. Kennedy, R. Eberhart, *Proceedings of IEEE International Conference on Neural Networks*, 2005, pp. 1942–1948.
- [28] S. Das, A. Abraham, A. Konar, *Stud. Comput. Intell.* 116 (2008) 1–38.
- [29] R. Storn, K. Price, *J. Global Optim.* 11 (1997) 341–359.
- [30] M. Srinivas, G.P. Rangaiah, *Comp. Chem. Eng.* 31 (2007) 760–772.
- [31] J.E. Dennis Jr., R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
- [32] J.A. Lopez, V.M. Trejos, C.A. Cardona, *Fluid Phase Equilib.* 248 (2006) 147–157.
- [33] J. Gmehling, U. Onken, W. Arlt, *Vapor–liquid Equilibrium Data Collection*, Chemistry Data Series, vol. I, Parts 1–8, DEHEMA, 1977–1990.
- [34] Y.S. Teh, G.P. Rangaiah, *Comp. Chem. Eng.* 27 (2003) 1665–1679.