Short communication

An improved ant colony optimization method and its application for the thermodynamic modeling of phase equilibrium

Jorge Adán Fernández-Vargas [a,b], Adrián Bonilla-Petriciolet [a,*],
Juan Gabriel Segovia-Hernández [b]

[a] Instituto Tecnológico de Aguascalientes, Aguascalientes 20256, Mexico
[b] Universidad de Guanajuato, Guanajuato 36050, Mexico

ARTICLE INFO

ABSTRACT

Ant colony optimization (ACO) is a stochastic optimization method that emulates the indirect communication among the individuals of an ant colony. This multi-agent strategy has been used to solve optimization problems from engineering applications, especially for combinatorial problems. Although ACO seems to be effective for global optimization, only few ACO algorithms have been reported for solving engineering applications problems with continuous decision variables. In this study, a new continuous ACO algorithm, with feasible region selection, has been implemented and applied to perform thermodynamic calculations related to the modeling of phase equilibrium. In particular, the numerical capabilities of this ACO algorithm have been tested in the global optimization of thermodynamic functions from parameter estimation, phase equilibrium and phase stability problems. These thermodynamic calculations are relevant for chemical engineering process simulators. However, they are classed as challenging nonconvex optimization problems with continuous decision variables. The applicability and effectiveness of the proposed ACO algorithm, namely ACOFRS, have been studied using a challenging set of thermodynamic benchmark problems with both liquid–liquid and vapor–liquid equilibrium. Results show that the proposed ACOFRS is an alternative method for performing global optimization in phase equilibrium calculations of multicomponent systems. In particular, ACOFRS is more robust for solving VLE parameter estimation problems and it outperformed other stochastic optimization methods such as Particle Swarm Optimization, Differential Evolution and Genetic Algorithms.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Computer-aided global optimization tools play an important role for process systems engineering in a broad diversity of chemical engineering applications. Particularly, the development of effective computational tools for phase equilibrium modeling has been recognized as a fundamental research topic in the literature of applied Thermodynamics. In fact, these calculations are necessary for the proper design, simulation and control of separations schemes involved in chemical and petrochemical processes. For example, reliable and precise thermodynamic calculations are essential during the analysis of phase behavior of multicomponent systems involved in distillation, absorption and other separation processes. Studies have estimated that 70–90% of process system costs (e.g., energy, operation and equipment) are usually related

to purification and separation units, where the thermodynamics of phase equilibrium influences their performance [1]. Formerly, phase equilibrium calculations (PEC) are performed to determine the thermodynamic identity (i.e., quantities and compositions) of phases at equilibrium, while the thermodynamic stability of the system is established using a phase stability analysis (PSA). These thermodynamic computations correspond to global optimization problems using the tangent plane distance function (TPDF) for PSA and the Gibbs free energy function (GF) for PEC [2]. On the other hand, parameter estimation (PaE) is crucial for developing thermodynamic models that can represent satisfactorily the phase behavior of mixtures involved in separation operations. In PaE problem, the parameters of thermodynamic models are used as optimization variables to obtain the best fit to measured experimental data. The objective functions commonly used for performing parameter estimation in phase equilibrium modeling are the least squares approach (PaE-LS) and the error-in-variable formulation (PaE-EIV). These non-convex problems also require global optimization tools to determine the global optimum parameters of

* Corresponding author. Tel.: +52 4499105002; fax: +52 4499105002.
E-mail address: petriciolet@hotmail.com (A. Bonilla-Petriciolet).

thermodynamic models for a proper representing of the thermodynamic data of phase behavior [2].

There are important numerical challenges (e.g., the presence of trivial solutions and several local optima) for performing the global optimization on PEC, PSA and PaE problems, most of them derived from the fact that thermodynamic models are non-linear and the objective functions are nonconvex [2]. In addition, the complexity of optimization problem increases when the system has several components and shows multi-phase behavior because the number of optimization variables depends on these problem characteristics. Therefore, the traditional numerical approaches used to solve these equilibrium problems may fail to find the global minimum solution.

Various optimization methods have been used to resolve PaE, PEC and PSA problems [2–23]. Several studies have documented that deterministic methods are theoretically convergent to the global optimum but the problem reformulation is usually necessary because these methods are restricted to satisfy specific properties of the objective function [2]. Besides, the computer time is usually high in multivariable problems. In contrast, stochastic global optimization (SGO) techniques do not require assumptions or mathematical transformations of the optimization problems because they work as black-box methods and, consequently, they are simple to implement and use [2]. Literature on PEC, PSA and PaE has shown that SGO methods offer a reasonable compromise between solution quality and numerical effort [2,3,6–8,12,14,16–23]. SGO methods applied for phase equilibrium calculations and modeling include traditional random search techniques (e.g., Simulated Annealing, Luus–Jaakola method), population-based and evolutionary algorithms (e.g., Genetic Algorithms, Swarm Optimization, Differential Evolution) and hybrid strategies (e.g., Differential Evolution with Tabu List, Random Tunneling Algorithm). Even though different SGO techniques have been applied on thermodynamic modeling, there are still specific aspects to be improved on them for solving reliably global optimization problems. For example, most of the meta-heuristics may fail to find the global optimum at early iterations, especially, when the phase behavior of multicomponent systems is analyzed. Therefore, it is important to develop and study novel SGO methods with the aim of identifying a better strategy for performing these chemical engineering calculations.

An attractive alternative for the robust resolution of multicomponent thermodynamic problems is the use of ant colony optimization (ACO) [24,25]. ACO algorithms have been successful to resolve engineering and real-life optimization problems. Specifically, this method has shown a remarkable performance for solving complex combinatorial problems including the traveling salesman and quadratic assignment problems [25–29]. In the case of real-life application problems, this meta-heuristic has been used for vehicle routing [30,31], scheduling problems [32], and dynamic optimization of chemical processes [33]. The application of ACO in thermodynamic calculations for the modeling of phase equilibrium behavior has not been reported. Furthermore, based on the fact that ACO was proposed as a combinatorial optimization tool, few studies have reported its application for solving engineering optimization problems with continuous decision variables [34–36].

This study introduces an improved ACO method for performing phase equilibrium calculations in multicomponent systems. In particular, the concept of the feasible region selection has been introduced to improve the performance of ACO in continuous global optimization. Our preliminary studies have shown that the proposed ACO method may outperform other ACO-type methods for solving large-scale global optimization problems with continuous decision variables. These results have motivated the extension and application of this novel ACO

method for performing thermodynamic calculations in multicomponent systems, which are relevant for process modeling and design in chemical engineering. Reliability and capabilities of this ACO-based method have been analyzed and discussed using a number of PaE, PEC and PSA problems. This note reports the first application of an ACO-type optimization method in thermodynamic calculations related to phase equilibrium modeling.

2. Ant colony optimization

2.1. Principles

Dorigo [24] proposed to emulate the collective behavior of real ants for solving combinatorial optimization problems. The conceptual inspiration of this metaheuristic is based on the behavior of ants for searching food and finding their way back to the nest using the chemical trail called *pheromone*. Pheromone trail guides the other ants towards the target point and one ant selects its path based on the pheromone quantity. The collective behavior that results from this stage is a *positive feedback* response and, consequently, more ants will follow a specific path to the food source [25]. The meta-heuristic ACO consists of three scheduled activities: the solution construction based on the ant performance, the update of pheromone trail and the daemon actions [26]. These activities are repeated during the search for the optimum solution. ACO was developed to resolve combinatorial optimization problems and the solution construction has been traditionally handled as a selecting procedure. For example, in the simple ACO algorithm (SACO), the selection is performed according to the *transition probability* [29]

$$p_{ij}^k(t) = \begin{cases} \dfrac{\tau_{ij}^\alpha(t)}{\sum_{j \in N_i^k} \tau_{ij}^\alpha(t)} & \text{if } j \in N_i^k \\ 0 & \text{if } j \notin N_i^k \end{cases} \tag{1}$$

where N_i^k is the set of feasible solution components (associated to i component) for the ant k at time t, $\tau_{ij}(t)$ is the pheromone trail of path selection ij and α is a positive value that determines the weight of pheromone information during the search. Pheromone update consists of two stages: intensification and evaporation. The first process is performed when every ant deposits a pheromone amount ($\Delta \tau^k(t)$) in every transition ij that has been chosen. Then, pheromone concentration is intensified considering this pheromone amount [29]

$$\tau_{ij}(t+1) \leftarrow \tau_{ij}(t) + \Delta \tau^k(t) \tag{2}$$

Most of the ACO algorithms simulate the natural process of pheromone evaporation using the next approach

$$\tau \leftarrow (1 - \rho) \cdot \tau \tag{3}$$

where $\rho \in [0, 1]$ is the evaporation rate of pheromone. This parameter causes that ants can "forget" the previous decisions performed to determine the path to the food source. Finally, the daemon actions improve the algorithm performance via the implementation of centralized actions, which are not realized by single ants [26]. ACO is based on simple actions, however, a complex collective behavior emerges as result of the indirect communication via pheromone trails, which is a process with a large random component.

Bilchev and Parmee [37] introduced the first application of ant colony optimization for objective functions with continuous decision variables, where only local search procedures were used. This ACO adaptation was extended and improved using the classification of both global and local ants to resolve continuous optimization problems [32–34,38]. In this algorithm, namely CACO, the search space is randomly explored by creating regions, where each region

represents a point with its corresponding fitness. Ants create new regions using evolutionary strategies and the solution is improved by the selection of regions with better fitness in the local step using the pheromone information. In other study [35], an ant colony approach with continuous interacting has been proposed, where ants may interact between them using the pheromone information and a direct communication. Monmarché et al. [39] proposed other ACO approach for continuous optimization using a global exploration that is based on a tandem running where one ant leads another one, and so on. Finally, Socha and Dorigo [36] proposed the ACO$_\mathbb{R}$ method for solving optimization problems with continuous decision variables. In ACO$_\mathbb{R}$, the ants use a dynamic probability distribution at each construction step. The description of the ACO algorithm proposed in this study is given in the following section.

2.2. ACO with feasible region selection (ACOFRS)

For continuous search spaces, an unconstrained global optimization problem can be stated as

$$\text{Minimize} f(x) \tag{4}$$

with respect to n_{var} decision variables. Note that x represents a feasible point in the solution space with determined dimensional components: $x^j (j = 1, 2, \ldots, n_{var})$. Values of the decision variables are restricted to the upper and lower bounds: $(x_{\min}^j \leq x^j \leq x_{\max}^j)$.

In this study, an improved ACO method, namely ACO with *feasible region selection*, has been proposed to enhance the convergence properties in continuous global optimization problems especially for several decision variables. The transition probability given by Eq. (1), the basic concepts of SACO algorithm and the regions scheme of CACO were used as the platform to define the general structure for the algorithm ACOFRS. Briefly, ACOFRS explores the solution space, just as the ants search their food, by analyzing different regions and, after some time, they are able to find the shortest path. First step of ACOFRS is the creation of a region file **r**. The number of regions (NR) is defined and each region represents a point $r_i (i = 1, \ldots, NR)$ that corresponds to a feasible solution of the optimization problem. For all n_{var} decision variables, a pheromone amount is deposited in each component $r_i^j (j = 1, \ldots, D)$ of these regions. A $NR \times n_{var}$ matrix $\boldsymbol{\tau}$ is used to store the pheromone trial with a typical element τ_i^j, which indicates the pheromone level in the ith region for the jth decision variable. The pseudo code of the ACOFRS algorithm is given in Fig. 1. In this pseudo code, $U \in [A, B]$ is a real number and $UI \in [a, b]$ is an integer number, which are randomly generated inside the given bounds. The initialization of **r** is performed randomly and the best solution is identified and stored in a comparison vector $\mathbf{r}_B$.

As stated, the stage of ant-based solution construction is a key component of ACO. In the proposed algorithm, the dimensional components of each region have a selection probability p when the ant k explores the search space at the time t. Then, based on Eq. (1), the selection procedure implemented in ACOFRS is given by

$$p_{r_i^j}^k(t) = \begin{cases} \dfrac{\tau_i^j(t)}{\sum_{i \in N^k} \tau_i^j(t)} & \text{if } i \in N^k \\ 0 & \text{if } i \notin N^k \end{cases} \tag{5}$$

where N^k is the result of the *feasible region selection*, which is different for each ant k. N^k is a set of NSR randomly selected elements (with at least two different elements) from the matrix of regions **r**. Parameter NSR controls the influence of the pheromone information and its upper limit is NR. Once all selections have been performed, a set of dimensional elements $r_{sj}^{j,k}$ with $j = 1, \ldots, n_{var}$ is defined and a subset of regions $M^k \in N^k$ emerges. During the

stage of solution construction, each ant can take the precise value of the corresponding dimensional component, or it can perform a *path search* with a given probability $SP \in [0, 1]$ using the following operator

$$x^{j,k} = \begin{cases} r_{sj}^{j,k} + U[0, 1] * (r_\alpha^{j,k} - r_\beta^{j,k}) & \text{if } U \in [0, 1] < SP \\ r_{sj}^{j,k} & \text{else} \end{cases} \tag{6}$$

This *path search* process allows the movement of an ant from the selected component to a nearby location. In this way, ACOFRS can explore continuous domains of decision variables. When all the values of each decision variable have been defined, a new region x^k (i.e., trial point) has been generated. The objective function is evaluated for this trial point and it will replace an existing region CO only if the new function value $f(x^k)$ is less than $f(r_{sCO})$ where CO is a comparison region obtained from M^k. The pheromone intensification step is given by

$$\tau_{sj}^{j,k} \leftarrow \tau_{sj}^{j,k} + 1 \tag{7}$$

$$\tau_{sCO}^{j,k} \leftarrow \tau_{sj}^{j,k} \tag{8}$$

Particularly, this strategy allows an implicit evaluation of the pheromone deposits by the algorithm because $\Delta\tau_{\text{int}}$ is not proportional to the solution quality and all ants deposit the same pheromone quantity [40]. Preliminary calculations indicate that this pheromone intensification approach improves the convergence performance of ACO. The pheromone of all regions is evaporated after the exploration of all NA ants using a similar mechanism to Eq. (7). In this study, NA was defined equal to NR and the intensification process of ACOFRS is controlled via the number of sub-regions $M^k \in N^k$. Finally, the stages of pseudo code given by Fig. 1 are performed until the stopping criterion is satisfied.

3. Description of thermodynamic problems used for testing the performance of ACOFRS

To evaluate the numerical performance of ACOFRS, several PaE, PEC and PSA examples have been selected from the chemical engineering literature [3,8,10–12,16–23], which have been used in other studies for analyzing and comparing global optimization tools. These optimization problems are characterized by the non-convexity of the objective function and involve continuous decision variables with vapor–liquid (VLE) or liquid–liquid (LLE) equilibrium. Tables 1–4 provide the objective functions, problem characteristics (i.e., the global minimum and the corresponding values of the decision variables) of all examples used in this study. A brief description of these thermodynamic calculations is given in the following subsections.

Phase equilibrium calculations. The global minimization of the Gibbs free energy (g) is the traditional thermodynamic framework for performing PEC. The main objective of PEC is to identify the types (i.e., vapor, liquid or solid) of phases co-existing at equilibrium as well as their quantity and compositions for a mixture of substances with feed composition z at a given pressure P and temperature T [2]. This optimization problem is subject to the mass balance equality constraints. We have used an unconstrained formulation for the global minimization of the Gibbs free energy; see Table 1. The selected PEC problems include multicomponent mixtures with liquid–liquid (LLE) and vapor–liquid equilibrium (VLE).

Phase stability analysis. This thermodynamic calculation is necessary to assess if the thermodynamic state of a mixture corresponds to the global minimum of the Gibbs free energy. The tangent plane distance function (i.e., TPDF) is the objective function for this

Table 1
Formulation of phase stability, phase equilibrium and parameter estimation problems.

Thermodynamic problem	Objective function	Decision variables	Optimization problem
Phase stability analysis	Tangent plane distance function (TPDF): $$TPDF = \sum_{i=1}^{c} y_i(\mu_i\|_y - \mu_i\|_z)$$ where c is the number of components of the mixture, $\mu_i\|_y$ and $\mu_i\|_z$ are the chemical potentials calculated at trial composition y and feed composition z, respectively	The decision variables are $\beta_i \in (0,1)$ using the following relationships: $n_{iy} = \beta_i z_i n_F \quad i = 1, ..., c$ $y_i = \frac{n_{iy}}{\sum_{j=1}^{c} n_{iy}} \quad i = 1, ..., c$ where n_{iy} are the mole numbers of component i in phase y and n_F is the total moles in the mixture under analysis	$\min_{\beta} \quad TPDF \quad 0 \le \beta_i \le 1 \quad i = 1, ..., c$
Phase equilibrium calculations	Gibbs free energy of mixing (g) $$g = \sum_{j=1}^{\pi}\sum_{i=1}^{c} n_{ij}\ln(x_{ij}\gamma_{ij}) = \sum_{j=1}^{\pi}\sum_{i=1}^{c} n_{ij}\ \ln\left(\frac{x_{ij}\hat{\phi}_{ij}}{\phi_i}\right)$$ where π is the number of phases at equilibrium and θ_{ij} denotes the composition (i.e., x or n) or thermodynamic property of component i in phase j	The decision variables are $\beta_{ij} \in (0,1)$ using the following relationships: $n_{i1} = \beta_{i1} z_i n_F \quad i = 1, ..., c$ $n_{ij} = \beta_{ij}\left(z_i n_F - \sum_{m=1}^{j-1} n_{im}\right) \quad i = 1, ..., c \quad j = 2, ..., \pi - 1$ $n_{i\pi} = z_i n_F - \sum_{m=1}^{\pi-1} n_{im} \quad i = 1, ..., c$	$\min_{\beta} \quad g \quad 0 \le \beta_{ij} \le 1 \quad i = 1, ..., c \quad j = 1, ..., \pi - 1$
VLE parameter estimation	Classical least squares criterion based on activity coefficients: $$F_{LS} = \sum_{j=1}^{ne}\sum_{i=1}^{c} \left(\frac{\gamma_{ij}^{exp} - \gamma_{ij}^{calc}}{\gamma_{ij}^{exp}}\right)^2$$ where γ_{ij}^{exp} is the experimental value for the activity coefficient of component i in jth experiment, γ_{ij}^{calc} is the calculated value for the activity coefficient of component i in jth experiment, and c is the number of components in the mixture	The decision variables of this parameter estimation problem are the $npar$ parameters $\theta = (\theta_1, ..., \theta_{npar})^T$ of thermodynamic model.	$\min_{\theta} \quad F_{LS} \quad \theta_{i,min} \le \theta_i \le \theta_{i,max} \quad i = 1, ..., npar$
VLE parameter estimation	Error-in-variable formulation: $F_{obj} =$ $$\sum_{j=1}^{ne}\sum_{i=1}^{c}\left[\frac{(x_{ij}^t - x_{ij})^2}{\sigma_{x_i}^2} + \frac{(y_{ij}^t - y_{ij})^2}{\sigma_{y_i}^2}\right] + \sum_{j=1}^{ne}\left[\frac{(T_j^t - T_j)^2}{\sigma_T^2} + \frac{(P_j^t - P_j)^2}{\sigma_P^2}\right]$$ ne is the number of state variables, z_{ij}^t is the unknown "true" value of ith state variable in jth measurement, and σ_i is the standard deviation associated with the measurement of ith state variable	The decision variables of EIV problem are the set of z_{ij}^t and the model parameters θ	$\min_{z^t,\theta} \quad F_{EIV}$ subject to $g(z_{ij}^t, \theta) = 0 \quad i = 1, ..., nest \quad j = 1, ..., ne$ $\theta_{i,min} \le \theta_i \le \theta_{i,max} \quad i = 1, ..., npar$ $z_{ij,min}^t \le z_{ij}^t \le z_{ij,max}^t \quad i = 1, ..., nest \quad j = 1, ..., ne$

Fig. 1. Pseudo code of ACOFRS method.

analysis [2]. This function is given by the distance between the tangent plane at the given phase composition and the Gibbs free energy surface. The global minimization of TPDF has been performed using an unconstrained approach according to the problem formulation defined in Table 1. Similar to PEC, selected PSA examples involve mixtures with LLE and VLE.

Parameter estimation. As stated in the introduction section, the parameter estimation for the modeling of phase equilibrium behavior of multicomponent mixtures implies a global optimization problem where the task is the determination of adjustable parameters of a non-linear thermodynamic model (e.g., EoS or local composition models) used to fit a given set of experimental data (e.g., equilibrium compositions, T and P). In Thermodynamics, the objective function used for parameter estimation is usually derived from the traditional least squares (PaE-LS) or the error-in-variables (PaE-EIV) formulations [2]. Parameters of thermodynamic models are determined in the first approach while,

in the second formulation, both data reconciliation and parameter estimation are performed. Both optimization problems can be defined using an unconstrained approach, see Table 1. PaE-LS and PaE-EIV examples include VLE data of binary systems and local composition models have been used for testing the performance of ACOFRS in this type of thermodynamic calculations.

4. Results

4.1. Implementation of ACOFRS

Fortran® codes were developed for implementing the optimization algorithm and all thermodynamic problems used in this study. Overall, 34 thermodynamic problems were considered: 9 PEC, 9 PSA and 16 PaE problems. These optimization problems are multivariable and non-convex. Following previous

Table 2
Description of PEC and PSA problems used for testing the performance of ACOFRS method.

No.	System	Feed conditions	Thermodynamic models and reference	Global optimum	
				g	TDPF
1	n-Butyl acetate + water	$n_F = (0.5, 0.5)$ at 298 K and 101.325 kPa	NRTL model with model parameters reported by [12]	−0.020198	−0.032466
2	Toluene + water + aniline	$n_F = (0.29989, 0.20006, 0.50005)$ at 298 K and 101.325 kPa	NRTL model with model parameters reported by [4]	−0.352957	−0.294540
3	$N_2 + C_1 + C_2$	$n_F = (0.3, 0.1, 0.6)$ at 270 K and 7600 kPa	SRK EoS with classical mixing rules and model parameters reported by [17]	−0.547791	−0.015767
4	$C_1 + H_2S$	$n_F = (0.9813, 0.0187)$ at 190 K and 4053 kPa	SRK EoS with classical mixing rules and parameters reported by [12]	−0.019892	−0.003932
5	$H_2O + CO_2 + $ 2-propanol + ethanol	$n_F = (0.99758, 0.00003, 0.00013, 0.00226)$ at 350 K and 2250 kPa	SRK EoS with classical mixing rules and model parameters reported by [11]	−0.0048272	−0.012650
6	$C_2 + C_3 + C_4 + C_5 + C_6$	$n_F = (0.401, 0.293, 0.199, 0.0707, 0.0363)$ at 390 K and 5583 kPa	SRK EoS with classical mixing rules and model parameters reported by [17]	−1.183653	−0.000002
7	$C_1 + C_2 + C_3 + C_4 + C_5 + C_6 + C_{7-16} + C_{17+}$	$n_F = (0.7212, 0.09205, 0.04455, 0.03123, 0.01273, 0.01361, 0.07215, 0.01248)$ at 353 K and 38500 kPa	SRK EoS with classical mixing rules and model parameters reported by [11]	−0.838783	−0.002688
8	$C_1 + C_2 + C_3 + iC_4 + C_4 + iC_5 + C_5 + C_6 + iC_{15}$	$n_F = (0.614, 0.10259, 0.04985, 0.008989, 0.02116, 0.00722, 0.01187, 0.01435, 0.16998)$ at 314 K and 2010.288 kPa	SRK EoS with classical mixing rules and model parameters reported by [12]	−0.769772	−1.486205
9	$C_1 + C_2 + C_3 + C_4 + C_5 + C_6 + C_7 + C_8 + C_9 + C_{10}$	$n_F = (0.6436, 0.0752, 0.0474, 0.0412, 0.0297, 0.0138, 0.0303, 0.0371, 0.0415, 0.0402)$ at 435.35 K and 19150 kPa	SRK EoS with classical mixing rules and model parameters reported by [17]	−0.00002	−1.121176

studies on stochastic methods, e.g. [3,20–23], the performance analysis of ACOFRS is based on the results from resolving each thermodynamic problem 100 times with random initial values of decision variables and using a different random number seed. Two different local optimizers (i.e., Quasi-Newton method and Nelder–Mead method) have been used as intensification approach to improve the precision of solution found by ACOFRS. Quasi-Newton method (QNF) calculates the gradient via finite differences and approximates the Hessian matrix according to BFGS formula, while Nelder–Mead method (NMF) is a direct search

Table 3
Description of VLE data modeling problems with LS formulation used for testing the performance of ACOFRS method.

No.	System	Data	Thermodynamic models	Decision variables	Global optimum	Adjusted model parameters
1	Tert butanol + 1 butanol	$P = 100$ mm Hg $n_{dat} = 9$	Wilson and ideal gas model	$n_{var} = 2$ $\theta_1, \theta_2 \in (-8500, 320,000)$	0.01026	$\theta_1 = -567.96$ $\theta_2 = 745.33$
2		$P = 700$ mm Hg $n_{dat} = 9$			0.013690	$\theta_1 = -733.95$ $\theta_2 = 1318.23$
3		$P = 500$ mm Hg $n_{dat} = 9$			0.006852	$\theta_1 = -718.01$ $\theta_2 = 1264.74$
4	Water + 1,2 ethanediol	$P = 430$ mm Hg $n_{dat} = 18$	Wilson and ideal gas model	$n_{var} = 2$ $\theta_1, \theta_2 \in (-8500, 320,000)$	1.039134	$\theta_1 = 5072.361$ $\theta_2 = -1921.62$
5			UNIQUAC and ideal gas model	$n_{var} = 2$ $\theta_1, \theta_2 \in (-5000, 20,000)$	1.408547	$\theta_1 = -1131.84$ $\theta_2 = 3617.65$
6			NRTL and ideal gas model	$n_{var} = 3$ $\theta_1, \theta_2 \in (-2000, 5000)$ $\alpha_{12} \in (0.01, 10.0)$	1.253531	$\theta_1 = -678.99$ $\theta_2 = 3046.13$ $\alpha_{12} = 0.621375$
7	Benzene + hexafluorobenzene	$T = 50\,^\circ$C $n_{dat} = 11$	Wilson and ideal gas model	$n_{var} = 2$ $\theta_1, \theta_2 \in (-8500, 320,000)$	0.008935	$\theta_1 = -424.08$ $\theta_2 = 983.06$
8		$P = 300$ mm Hg $n_{dat} = 17$			0.014860	$\theta_1 = -432.49$ $\theta_2 = 992.85$
9		$P = 760$ mm Hg $n_{dat} = 29$			0.014616	$\theta_1 = -334.70$ $\theta_2 = 704.74$
10		$T = 30\,^\circ$C $n_{dat} = 10$			0.011783	$\theta_1 = -467.76$ $\theta_2 = 1313.94$

Table 4
Description of VLE data modeling problems with EIV formulation used for testing the performance of ACOFRS method.

No.	System	Data	Thermodynamic models	Decision variables[a]	Global optimum	Adjusted model parameters
11	Benzene + hexafluorobenzene	$P = 500$ mmHg; $n_{dat} = 16$ σ (0.003,0.0029,1.7,0.83)	Wilson and ideal gas model	$n_{var} = 34$	19.998720	$\theta_1 = -429.85$ $\theta_2 = 1029.32$
12		$P = 300$ mmHg; $n_{dat} = 17$ σ (0.003,0.0029,1.7,0.083)		$n_{var} = 36$	42.343724	$\theta_1 = -437.72$ $\theta_2 = 1003.12$
13		$P = 30\,°C$; $n_{dat} = 10$ σ (0.001,0.01,0.75,0.1)		$n_{var} = 22$	11.898795	$\theta_1 = -472.00$ $\theta_2 = 1274.32$
14		$P = 40\,°C$; $n_{dat} = 10$ σ (0.001,0.01,0.75,0.1)		$n_{var} = 22$	11.170496	$\theta_1 = -462.51$ $\theta_2 = 1197.31$
15		$P = 50\,°C$; $n_{dat} = 11$ σ (0.003,0.0029,1.7,0.083)		$n_{var} = 24$	25.671042	$\theta_1 = -415.76$ $\theta_2 = 944.57$
16		$P = 60\,°C$; $n_{dat} = 10$ σ (0.003,0.0029,1.7,0.083)		$n_{var} = 22$	19.401593	$\theta_1 = -437.05$ $\theta_2 = 1065.04$

[a] For all EIV examples, the limits of decision variable are: $\theta_1, \theta_2 \in (-10,000, 200,000)$, $x_{ij}^t \in (x_{ij} - 3\sigma, x_{ij} + 3\sigma)$ and $T^t \in (T - 3\sigma, T + 3\sigma)$.

Table 5
Suggested values of ACOFRS parameters for solving thermodynamic calculations in phase equilibrium modeling.

Parameter	Studied values	Suggested values for		
		TPDF	g	Parameter estimation
SP	0.1–0.9	0.6	0.7	0.8
NSR	2–$5 * n_{var}$	$2 * n_{var}$	$3 * n_{var}$	$2 * n_{var}$

method that does not require derivatives of the objective function.

The performance of ACOFRS, with and without a local optimizer, has been analyzed using the success rate (SR, %) of finding the global optimum with a predetermined precision, and using a specific stopping condition. This success rate was calculated from 100 independent trials performed for each problem. A trial of ACOFRS is successful if the following condition holds: $|f^{ACOFRS} - f^*| \leq err$ where f^{ACOFRS} is the best function value obtained by ACOFRS and f^* is the known global minimum for the selected thermodynamic example. An error value $(err) = 10^{-05}$ was selected to avoid counting local minima as the global minima. Global success rate (GSR, %) for each type of thermodynamic calculation (i.e., PEC, PSA and PaE) has been calculated using

$$GSR = \sum_{i=1}^{n_{prob}} \left(\frac{SR_i}{n_{prob}} \right) \qquad (9)$$

where n_{prob} is the number of tested problems and SR_i is the individual success rate for each problem. All thermodynamic problems were solved using two stopping conditions for ACOFRS. First convergence criterion corresponds to the maximum number of iterations of ACOFRS ($Iter_{MAX}$), and second condition is an improvement-based stopping criterion defined as the maximum number of iterations (SC_{MAX}) without improvement in the best solution found by ACOFRS. For each thermodynamic problem, the mean number of function evaluations has been also reported. Parameters SP and NSR of ACOFRS were tuned for solving the different thermodynamic problems. Representative examples were selected for this purpose, and ACOFRS parameters were tuned one at a time. Suggested values of ACOFRS parameters for each thermodynamic calculation are given in Table 5.

4.2. Application of ACOFRS on PSA and PEC problems

Fig. 2 shows GSR of ACOFRS without the local optimization at different values of $Iter_{MAX}$ for all PSA and PEC problems. Particularly, the maximum GSR of ACOFRS is 61.6% in PEC problems and there is a natural tendency of improving the success for finding the global optimum solution with the increment of the number of iterations of ACOFRS. However, GSR of ACOFRS does not improve significantly after 500 iterations. For illustration, Fig. 3 shows the convergence profiles of ACOFRS without local optimization for selected thermodynamic problems: PSA problem No. 3 (T3) and PEC problem No. 7 (G7). It is interesting to remark that ACOFRS may show a good numerical performance for finding the global optimum solution in some examples with high precision in the solution vector with error values lower than $1E - 04$, see Fig. 3.

Performance results of ACOFRS in combination with QNF and NMF for both PSA and PEC are reported in Fig. 4. As expected, the implementation of these local optimizers improves the success rates of ACOFRS in all cases especially at lower iteration levels due to the capability of local optimizer for finding a high precision of the global solution in few function evaluations. For example, ACOFRS without the local optimization was able to find the global optimum of PEC examples G2 and G3 in all trials performed for $Iter_{MAX} \geq 750$ while ACOFRS + QNF showed a 100% reliability for
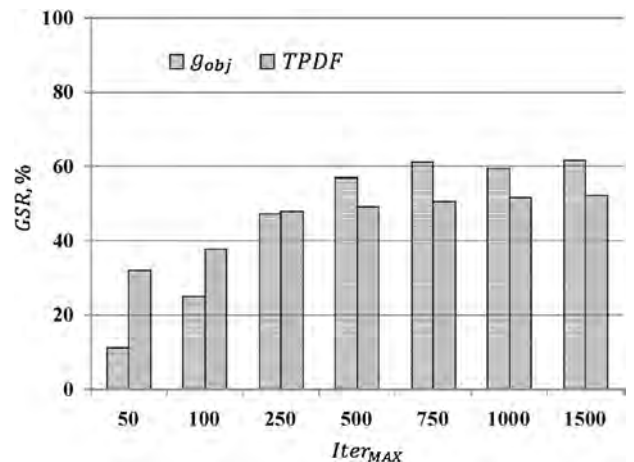

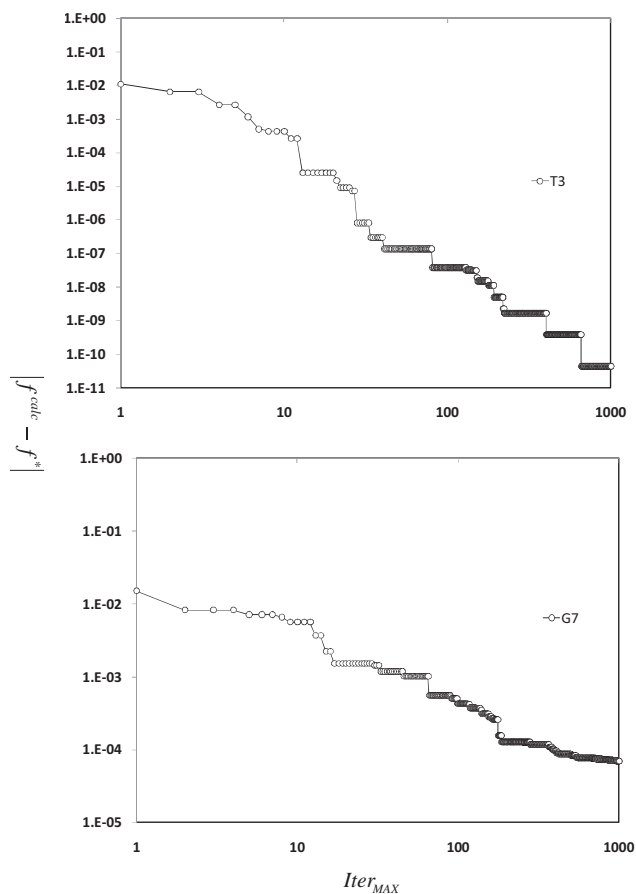**Fig. 2.** GSR versus $Iter_{MAX}$ of ACOFRS in PEC and PSA problems.

Fig. 3. Convergence profile of ACOFRS for selected thermodynamic problems.

solving the examples G2, G3, G7, G8 and G9 in the entire range of tested iteration levels.

In PSA problems, the maximum *GSR* of ACOFRS was 52.1% at $Iter_{MAX} > 250$. However, this success rate increased up to 72.7% with the use of NMF as intensification method (see Fig. 4a). For both PEC and PSA problems, ACOFRS + QNF showed a better performance than that obtained for ACOFRS + NMF at lower iteration levels. However, after 500 iterations ACOFRS + NMF may offer the best reliability for solving these thermodynamic problems. For example, ACOFRS + NMF can find the global optimum of PSA problems T1, T3, T4 and T8 with a 100% reliability. Overall, PSA problems T5 and T9 are the most challenging and ACOFRS failed several times to find the global optimum solution in the calculations performed. In particular, PSA problem T5 is very useful for testing and assessing the performance of novel global optimization methods for the modeling of phase equilibrium. Table 6 provides a summary of the performance metrics of ACOFRS + QNF for PEC and PSA problems using SC_{MAX} as stopping condition. It is clear that the numerical effort of ACOFRS increased with the value of SC_{MAX} but the increment of success rate is not substantial except for PEC problem G4 where $SR = 59\%$ at $SC_{MAX} = 10$ and $SR = 96\%$ at $SC_{MAX} = 50$, respectively. For selected thermodynamic problems, the stopping condition $SC_{MAX} = 25$ provides an adequate balance between reliability and efficiency of ACOFRS + QNF for solving these thermodynamic problems. Overall, ACOFRS showed a better reliability for solving PEC problems than that obtained for PSA problems.

4.3. Application of ACOFRS on PaE problems

Figs. 5–7 show the performance of ACOFRS for solving parameter estimation problems. It is clear that the numerical performance
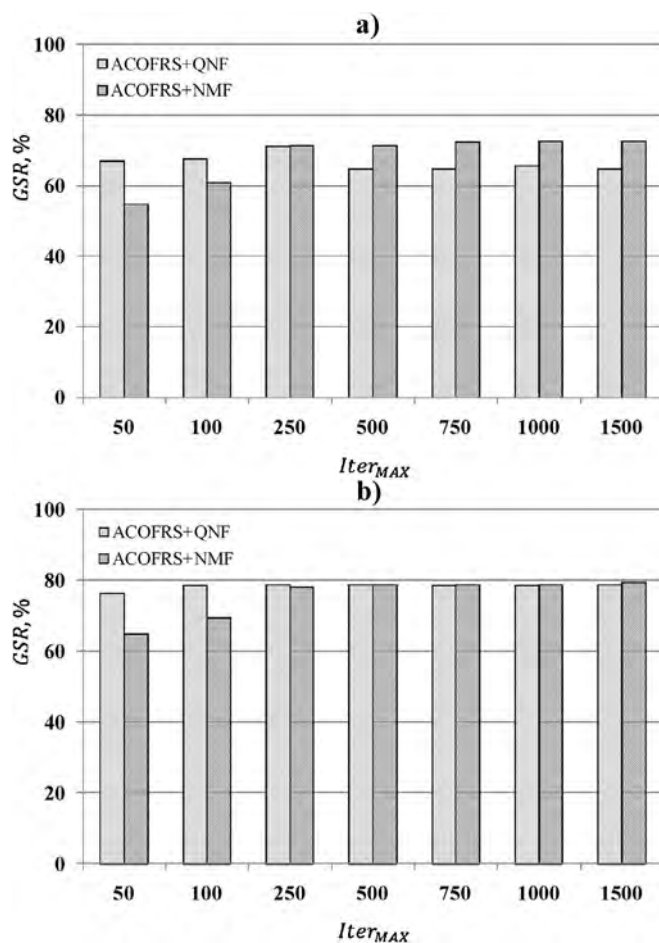
Fig. 4. *GSR* versus $Iter_{MAX}$ of ACOFRS with local optimizers QNF and NMF for (a) PSA and (b) PEC problems.

of the proposed method significantly depends on $Iter_{MAX}$, especially if ACOFRS is used without the local optimization for the intensification stage. Results show that a larger number of regions for ACOFRS improves the exploration capabilities of the stochastic method and increases the probability for finding the global optimum, see results reported in Figs. 5 and 6. For example, ACOFRS using $NA = NR = 50 * nvar$ is very reliable for solving PaE-LS problems P4 and P9. However, the numerical effort may increase considerably because more function evaluations are performed during

Fig. 5. *GSR* versus $Iter_{MAX}$ of ACOFRS for PaE problems using LS formulation.

Table 6
Numerical performance of ACOFRS + QNF for PSA and PEC problems using SC_{MAX} as stopping criterion.

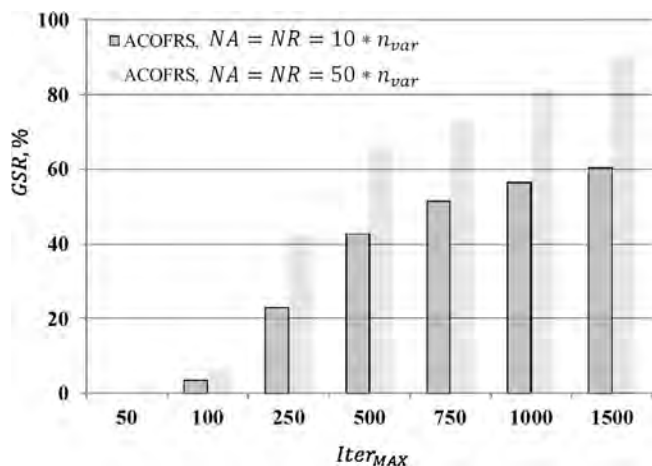| No.	NFE_{Total}			SR (%)		
	$SC_{MAX}=10$	$SC_{MAX}=25$	$SC_{MAX}=50$	$SC_{MAX}=10$	$SC_{MAX}=25$	$SC_{MAX}=50$
Numerical performance of ACOFRS for phase stability analysis[a]						
1	1120	1758	3190	86	95	95
2	2218	4793	8682	99	99	96
3	911	2532	5514	100	100	100
4	512	1222	2329	100	100	100
5	-	-	-	-	-	-
6	1379	4163	11,324	57	65	65
7	2663	8052	25,977	32	48	52
8	4352	41,703	56,737	100	56	58
9	3240	13,749	46,130	28	29	25
	NFE_{Total}			SR (%)		
	$SC_{MAX}=10$	$SC_{MAX}=25$	$SC_{MAX}=10$	$SC_{MAX}=25$	$SC_{MAX}=10$	$SC_{MAX}=25$
Numerical performance of ACOFRS for phase equilibrium calculations[a]						
1	460	2201	5257	82	93	97
2	599	2133	7549	96	100	100
3	699	1967	5040	100	100	100
4	380	1028	2455	59	89	96
5	-	-	-	-	-	-
6	1305	3872	8187	10	15	17
7	2109	6087	15,934	100	100	100
8	2324	20,789	66,401	100	100	100
9	2650	9778	23,618	100	99	100

[a] Symbol '-' indicates that NFE_{Total} is not reported since stochastic method showed 0% SR.
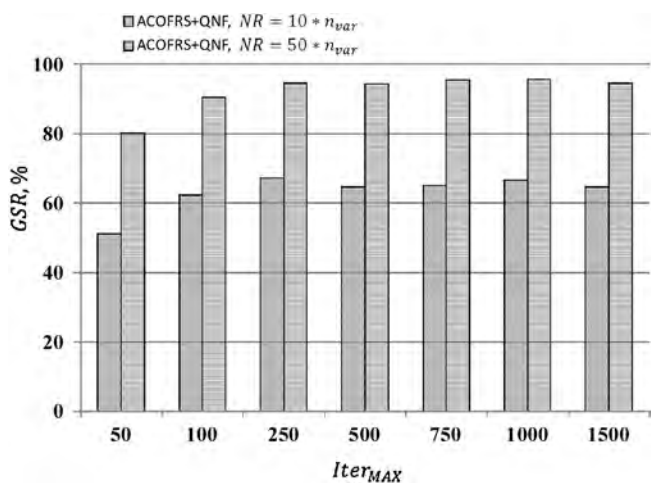
Fig. 6. GSR versus $Iter_{MAX}$ of ACOFRS + QNF in PaE problems using LS formulation.
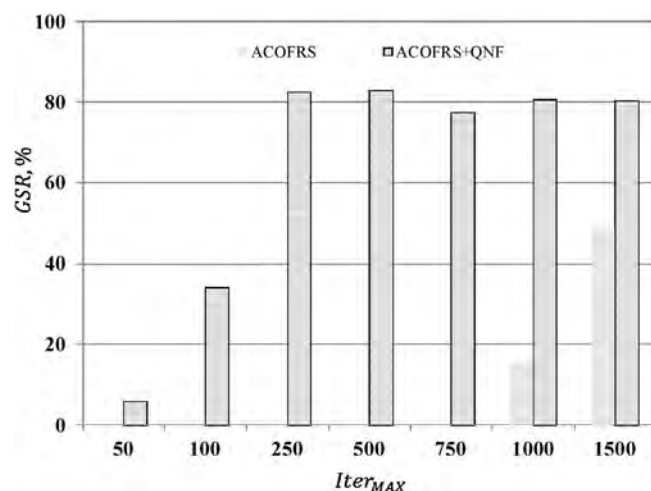
Fig. 7. GSR versus $Iter_{MAX}$ of ACOFRS with and without QNF in PaE problems using EIV formulation.

the search of the global solution. Note that the function evaluations (i.e., numerical effort) are proportion to the number of regions used in ACOFRS. For PaE-LS formulation, ACOFRS + QNF showed SR from 56 to 100% for selected examples. The numerical behavior of ACOFRS, with and without the local optimization, for solving PaE-EIV problems is reported in Fig. 7. Maximum GSR is 82.8% for PaE-EIV examples P11–P16 using ACOFRS + QNF. It is conve-

nient to remark that previous studies [21] have reported GSR of 66 and 63% for these examples using Differential Evolution and Differential Evolution with Tabu List. These stochastic methods have been recognized as promising metaheuristics for solving global optimization problems in the context of chemical engineering. It is clear that ACOFRS may offer a better numerical performance

Table 7
Numerical performance of ACOFRS + QNF for selected PaE problems using $Iter_{MAX}$ as stopping criterion.

No.	n_{var}	NFE_{Total} for $Iter_{MAX}$						
		50	100	250	500	750	1000	1500
1	2	1055	2055	5049	10,042	15,036	20,033	30,025
6	3	1577	3075	7565	15,051	22,549	30,047	45,039
11	34	18,024	34,642	85,670	170,639	255,668	340,625	510,610
12	36	18,797	36,705	90,705	180,691	270,681	360,667	540,639
13	22	13,061	24,917	60,916	114,465	168,148	221,532	330,811
15	24	16,701	27,757	63,738	120,639	180,596	240,540	360,387

Table 8
Numerical performance of ACOFRS and ACOFRS + QNF for PaE problems using SC_{MAX} as stopping criterion.

No.	SC_{MAX}	NFE		NFE_{Total}	SR (%)	
		ACOFRS	QNF		ACOFRS	ACOFRS + QNF
1	$6 * n_{var}$	1372	1833	3205	0	38
	$12 * n_{var}$	3401	31	3432	0	53
2	$6 * n_{var}$	1323	27	1350	0	31
	$12 * n_{var}$	3292	30	3322	1	59
3	$6 * n_{var}$	1244	478	1721	0	38
	$12 * n_{var}$	2994	28	3023	0	53
4	$6 * n_{var}$	1095	615	1710	6	39
	$12 * n_{var}$	3182	7	3189	48	65
5	$6 * n_{var}$	1095	17	1112	0	36
	$12 * n_{var}$	2606	11	2617	10	47
6	$6 * n_{var}$	1989	72	2061	1	56
	$12 * n_{var}$	6997	36	7033	5	56
7	$6 * n_{var}$	1173	41	1213	0	39
	$12 * n_{var}$	3281	25	3306	13	60
8	$6 * n_{var}$	1076	1234	2309	0	35
	$12 * n_{var}$	2928	25	2953	10	60
9	$6 * n_{var}$	1155	68	1223	0	56
	$12 * n_{var}$	2692	43	2735	0	90
10	$6 * n_{var}$	1174	71	1245	2	44
	$12 * n_{var}$	3110	24	3134	36	61
11	$6 * n_{var}$	510340	270	510610	0	97
	$12 * n_{var}$	510340	270	510610	0	97
12	$6 * n_{var}$	536393	279	536672	0	89
	$12 * n_{var}$	540360	279	540639	0	87
13	$6 * n_{var}$	327774	424	328198	79	80
	$12 * n_{var}$	325087	569	325657	72	73
14	$6 * n_{var}$	304262	999	305261	66	66
	$12 * n_{var}$	327969	813	328782	70	70
15	$6 * n_{var}$	354216	147	354363	77	79
	$12 * n_{var}$	357720	166	357886	71	72
16	$6 * n_{var}$	316151	141	316292	79	79
	$12 * n_{var}$	325334	151	325485	75	76

for solving the parameter estimation problems using error-in-variable formulation than those obtained with other stochastic optimization methods. This result indicated that the diversification approach used in ACOFRS is more effective for exploring the objective functions of error-in-variable problems where several decision variables are involved. In general, the performance of ACOFRS for finding a precise solution of parameter estimation problems increased with the use of the local optimization method, especially at early iterations. In fact, the use of local optimizer for the intensification stage in stochastic optimization methods is more relevant when large-scale optimization problems are resolved. Finally, Tables 7 and 8 summarize the performance metrics of ACOFRS using both stopping conditions $Iter_{MAX}$ and SC_{MAX} for solving parameter estimation problems.

In summary, the proposed ACOFRS method is an alternative optimization tool for solving thermodynamic problems related to phase equilibrium modeling and calculations involved in process design. Bases on the results reported in other studies [22,23], ACOFRS may compete with other stochastic optimization methods such as Covariant Matrix Adaptation Evolution Strategy [22] or Modified Cuckoo Search [23] for performing phase stability and equilibrium calculations. However, ACOFRS outperforms the optimization capabilities of Differential Evolution, Particle Swarm Optimization, Genetic Algorithms, Simulated Annealing, Harmony Search and other hybrid stochastic methods [21,41,42] for solving VLE parameter estimation problems. According to the No Free

Lunch Theorem for optimization [43], a general-purpose universal optimization method is theoretically impossible and the only way one strategy can outperform another if it is specialized to the specific problem under consideration. This implies that there are no formal assurances that a specific stochastic method used for solving several types of thermodynamic calculations will be at all effective (i.e., there is no an universally successful stochastic optimization method). Therefore, it is expected that ACOFRS may show different numerical performance in solving several types of thermodynamic calculations and it could be the best approach for performing one or more types of thermodynamic calculations in comparison to the results obtained with other stochastic optimization methods.

5. Conclusions

In this study, we have introduced an improved ant colony optimization method for performing thermodynamic calculations related to phase equilibrium modeling. The proposed ACOFRS was capable to solve most of the PEC and PS problems with a good performance. But, ACOFRS have shown outstanding results for solving parameter estimation problems using both the least squares and the error-in-variable formulations. Therefore, we suggest the application of this ACO-type method for solving parameter estimation problems of multicomponent systems. In further studies, we will

seek to improve the convergence properties of this method at early iterations with the aim of developing a more robust and effective stochastic global optimization strategy. Note that this aspect remains an open problem in the research of numerical performance of stochastic optimization methods.

References

[1] W.A. Wakeman, R.P. Stateva, Rev. Chem. Eng. 20 (2004) 1–56.
[2] H. Zhang, A. Bonilla-Petriciolet, G.P. Rangaiah, Open Thermodyn. J. 5 (2010) 71–92.
[3] M. Srinivas, G.P. Rangaiah, Comp. Chem. Eng. 31 (2007) 760–772.
[4] C.M. McDonald, C.A. Floudas, Comp. Chem. Eng. 19 (1995) 1111–1141.
[5] A.C. Sun, W.D. Seider, Fluid Phase Equilib. 103 (1995) 213–249.
[6] Y.P. Lee, G.P. Rangaiah, R. Luus, Comp. Chem. Eng. 23 (1999) 1183–1191.
[7] Y. Zhu, H. Wen, Z. Xu, Chem. Eng. Sci. 55 (2000) 3451–3459.
[8] A.L.H. Costa, F.P.T. da Silva, F.L.P. Pessoa, Braz. J. Chem. Eng. 17 (2000) 349–353.
[9] F. Jalali, J.D. Seader, Comp. Chem. Eng. 24 (2000) 1997–2008.
[10] C.Y. Gau, J.F. Brennecke, M.A. Stadtherr, Fluid Phase Equilib. 168 (2000) 1–18.
[11] S.T. Harding, C.A. Floudas, AIChE J. 46 (2000) 1422–1440.
[12] G.P. Rangaiah, Fluid Phase Equilib. 187/188 (2001) 83–109.
[13] D.V. Nichita, S. Gomez, E. Luna, Comp. Chem. Eng. 24 (2002) 1703–1724.
[14] Y.S. Teh, G.P. Rangaiah, Comp. Chem. Eng. 27 (2003) 1665–1679.
[15] G.I. Burgos-Solórzano, J.F. Brennecke, M.A. Stadtherr, Fluid Phase Equilib. 219 (2004) 245–255.
[16] M. Srinivas, G.P. Rangaiah, Comp. Chem. Eng. 24 (2006) 1400–1415.
[17] A. Bonilla-Petriciolet, R. Vázquez-Román, G.A. Iglesias-Silva, K.R. Hall, Ind. Eng. Chem. Res. 45 (2006) 4764–4772.
[18] A. Bonilla-Petriciolet, U.I. Bravo-Sanchez, F. Castillo-Borja, J.G. Zapiain-Salinas, J.J. Soto-Bernal, Braz. J. Chem. Eng. 24 (2007) 151–162.
[19] V.H. Álvarez, R. Larico, Y. Ianos, M. Aznar, Braz. J. Chem. Eng. 25 (2008) 409–418.
[20] A. Bonilla-Petriciolet, J.G. Segovia-Hernández, Fluid Phase Equilib. 289 (2010) 110–121.
[21] A. Bonilla-Petriciolet, G.P. Rangaiah, J.G. Segovia-Hernández, Fluid Phase Equilib. 287 (2010) 111–125.
[22] S.E. Fateen, A. Bonilla-Petriciolet, G.P. Rangaiah, Chem. Eng. Res. Des. 90 (2012) 2051–2071.
[23] V. Bhargava, S.E.K. Fateen, A. Bonilla-Petriciolet, Fluid Phase Equilib. 337 (2013) 191–200.
[24] M. Dorigo, PhD thesis, Politecnico di Milano, Italy, 1992.
[25] M. Dorigo, V. Maniezzo, A. Colorni, Proc. IEEE Trans. Syst. Man Cybern. 26 (1996) 29–41.
[26] M. Dorigo, G. Di Caro, L.M. Gambardella, Artif. Life 5 (1999) 137–172.
[27] T. Stützle, M. Dorigo, in: D. Corne, M. Dorigo, F. Glover (Eds.), New Ideas in Optimization, McGraw-Hill, New York, 1999.
[28] L.M. Gambardella, E. Taillard, M. Dorigo, J. Oper. Res. Soc. 50 (1999) 167–176.
[29] M. Dorigo, T. Stützle, Proceedings of International Conference on Evolutionary Computation, 2001, pp. 253–258.
[30] M. Reimann, K. Doerner, R. Hartl, D-ants: savings based ants divide and conquer the vehicle routing problems, Comput. Oper. Res. 31 (2004) (2004) 563–591.
[31] B. Yu, Z.Z. Yang, B. Yao, Eur. J. Oper. Res. 196 (2009) 171–176.
[32] V.K. Jayaraman, B.D. Kulkarni, S. Karale, P. Shelokar, Comp. Chem. Eng. 24 (2000) 1901–1912.
[33] J. Rajesh, K. Gupta, H.S. Kusumakari, V.K. Jayaraman, B.D. Kulkarni, Comput. Chem. 25 (2001) 559–583.
[34] M. Mathur, S.B.S.P. Karale, V.K. Jayaraman, B.D. Kulkarni, Ind. Eng. Chem. Res. 39 (2000) 3814–3822.
[35] J. Dréo, P. Siarry, Future Gen. Comput. Syst. 20 (2004) 841–856.
[36] K. Socha, M. Dorigo, Eur. J. Oper. Res. 185 (2008) 1155–1173.
[37] G. Bilchev, I.C. Parmee, Proceedings of Adaptive Computing in Engineering Design and Control, 1996, pp. 145–151.
[38] M. Wodrich, C. Bilchev, Control Cybern. 26 (1999) 413–446.
[39] N. Monmarché, G. Venturini, M. Slimane, Future Gen. Comput. Syst. 16 (2000) 937–946.
[40] A.P. Engelbrecht, Computational Intelligence an Introduction, John Wiley & Sons, England, 2007.
[41] H. Zhang, G.P. Rangaiah, A. Bonilla-Petriciolet, Ind. Eng. Chem. Res. 50 (2011) 10047–10061.
[42] A. Bonilla-Petriciolet, Fluid Phase Equilib. 332 (2012) 7–20.
[43] D.H. Wolpert, IEEE Trans. Evolut. Comput. 1 (1997) 67–82.