

PARAMETER ESTIMATION FOR VLE CALCULATION BY GLOBAL MINIMIZATION: THE GENETIC ALGORITHM

V. H. Alvarez¹, R. Larico², Y. Ianos² and M. Aznar^{1*}

¹School of Chemical Engineering, State University of Campinas, CEP: 13081-970, Campinas - SP, Brazil.
E-mail: maznar@feq.unicamp.br.

²School of Electrical Engineering and Computation, State University of Campinas,
P.O. Box 6066, CEP: 13081-970, Campinas - SP, Brazil.

(Received: February 1, 2007 ; Accepted: January 14, 2008)

Abstract - Vapor-liquid equilibrium calculations require global minimization of deviations in pressure and gas phase compositions. In this work, two versions of a stochastic global optimization technique, the genetic algorithm, the freeware MyGA program, and the modified mMyGA program, are evaluated and compared for vapor-liquid equilibrium problems. Reliable experimental data from the literature on vapor liquid equilibrium for water + formic acid, tert-butanol + 1-butanol and water + 1,2-ethanediol systems were correlated using the Wilson equation for activity coefficients, considering acid association in both liquid and vapor phases. The results show that the modified mMyGA is generally more accurate and reliable than the original MyGA. Next, the mMyGA program is applied to the CO₂ + ethanol and CO₂ + 1-n-butyl-3-methylimidazolium hexafluorophosphate systems, and the results show a good fit for the data.

Keywords: Genetic algorithm; Vapor-liquid equilibrium; Activity coefficient; Parameter estimation; Global optimization.

INTRODUCTION

Adjustable parameters are a common feature of most thermodynamic models for phase equilibrium calculations. The parameters for these models usually represent some physical property of the fluid, like molecular volume, shape and size, or intermolecular attractive forces. The interaction parameters can be temperature dependent or independent and their estimation is a common problem in many areas of science and engineering. The goal is to determine values of the model parameters that provide the best fit to measured data, generally based on some type of least squares or maximum likelihood criterion. This may require solution of a nonlinear and frequently nonconvex optimization problem, which often may be

formulated, either in constrained or unconstrained form.

Most of the existing methods for solving phase equilibrium and stability problems are local in nature and at best yield only local solutions. Use of global techniques in these problems is relatively unexplored and deserves greater investigation. Only deterministic global solvers, such as the homotopy continuation method (Sun and Seider, 1995), the interval Newton method (Hua et al., 1998a, 1998b), and the Lipschitz algorithm (Zhu and Xu, 1999), have been employed; all these methods use highly complex mathematics. On the other hand, stochastic optimization techniques use simple mathematics and have often been found to be as powerful and effective as deterministic methods in many engineering applications. There have been many

*To whom correspondence should be addressed

applications of genetic algorithms (GAs), originally proposed by Holland (1975), in chemical engineering (Ortega and Espiau, 2003) and in other fields. Genetic algorithms require only the objective function values and are highly likely to locate the global minimum.

In this work, the Wilson model is used to correlate experimental VLE data from water + formic acid, tert-butanol + butanol, and water + 1,2-ethanediol systems, where two versions of a genetic algorithm, the MyGa and the mMyGA programs, are studied and compared with the results of Gau et al. (2000).

PARAMETER ESTIMATION

Recently, alternative optimization methods have been proposed, among which neural nets and genetic algorithms (GAs) should be outlined. GA was first introduced by Holland (1975) and is adequately described in the literature (Goldberg, 1989; Mitchell, 1998).

GA is a stochastic global optimization technique that simulates natural evolution on the solution space of the optimization problems. GA is based on imitation of the evolutive behavior of a population of potential solutions (*chromosomes*) in each iteration (*generation*), where each individual competes with others to achieve the transmission of its genetic material to the next generation, taking into account the criterion of adaptability to the environment by combining some individuals of the current population according to predefined rules or operations; *crossover* and *mutation* operators are commonly used. Individuals are selected for reproduction based on their objective function values (fitness value) and the Darwinian principle of the survival of the fittest. GA is proven to yield better solutions along the evolving process, since good traits in the individuals of a generation are always passed to the next generation (Holland, 1975). Since there is no universal GA that would always produce the best performance for all optimization problems, one needs to perform preliminary tests with different versions of GA in order to identify the most suitable one for the problem(s) of interest.

An important step in a GA is the encoding of variables in a string structure (known as *genetic code*) to represent a point in the solution space. Three coding schemes – binary, gray, and real – are available. Since many engineering problems deal

with continuous variables, a real number is more appropriate, since it allows representation to the machine internal precision and also requires less memory (Haupt and Haupt, 1998). Michalewicz (1996) noted that real representation is more “natural” for continuous variables; thus a real-valued GA produces better results and is more efficient than its binary counterpart.

In this work, two versions of genetic algorithm, the MyGA program (MyGA, 2005) and the modified MyGA program (mMyGA), are evaluated and compared for vapor-liquid equilibrium problems. The MyGA program is the freeware version, written in Fortran, of a real-coded genetic algorithm developed by Yedder (2002), while the mMyGA is an adaptation (Alvarez, 2007).

Genetic Algorithm

Holland's genetic algorithm is intended to simulate nature's genetic algorithm in the following manner. The first step is to represent a solution to the problem by a string of genes that can take on some value from a specified finite range or alphabet. This string of genes, which represents a solution, is known as a chromosome. Then an initial population of chromosomes is constructed at random. In each generation, the fitness of each chromosome in the population is measured. The fittest chromosomes are then selected to produce offspring for the next generation, which inherit the best characteristics of both parents. After many generations of selection for the fittest chromosomes, the result is a population that is substantially fitter than the original. All genetic algorithms consist of the following main components:

a) Chromosomal Representation

Each chromosome represents a solution to the problem and is composed of a string of genes; for a multivariable equation, each variable is a gene and each possible solution is a chromosome. The binary alphabet [0,1] is often used to represent these genes, but sometimes, depending on the application, integers or real numbers are used.

b) Initial Population and Population

Once a suitable representation has been decided upon for the chromosomes, it is necessary to create an initial population to serve as the starting point for

the genetic algorithm. This initial population is usually created randomly. From empirical studies, over a wide range of function optimization problems, a population size of 10 times the number of variables is usually recommended (Deb, 2000). The population size is represented for N_{pop} .

c) Fitness Evaluation

Fitness evaluation involves defining an objective or fitness function against which each chromosome is tested for suitability for the environment under consideration. As the algorithm proceeds, the individual fitness of the "best" chromosome as well as the total fitness of the population as a whole is expected to increase.

d) Selection

Selection is based on the survival-of-the-fittest mechanism of nature. In order to mimic the survival of the fittest principle, individuals having better fitness values must have a higher chance of being selected. The selection procedure picks out two parent chromosomes, based on their fitness values, which are then used by the crossover and mutation operators (described below) to produce two offspring for the new population. The higher the fitness value, the higher the probability of the chromosome being selected for reproduction. Selection mechanisms, such as rank-based selection, elitism, niching, scaling and steady-state selection, have been proposed as operators. In particular, elitism determines that the best chromosome in each generation will pass unchanged to the new generation.

e) Crossover

Genetic operators used to create new individuals for the next population from selected individuals in the current population serve as search mechanisms in GA. After two parents have been selected by the selection method, crossover takes place. Crossover is an operator that mates the two parent chromosomes to produce two offspring that replace the parents. The crossover is carried out according to the crossover probability (p_c). A method for reducing the destructive effect of the crossover operator is called brood recombination (Banzhaf et al., 1998). In this procedure, the crossover is performed twice in order

to generate four offspring; the children are evaluated and sorted for fitness and the best two children are selected. The children replace the parents only if they have a better fitness; otherwise the parents remain.

f) Mutation

Mutation is a unary operator that creates a new solution by a random change in one or more of the chromosome's genes. The purpose of the mutation operator is to prevent the genetic population from converging to a local minimum and to introduce new possible solutions into the population. The mutation is carried out according to the mutation probability (p_m).

g) Termination

The termination method determines when the genetic process will stop evolving. There are several termination methods implemented in genetic algorithms. According to the *maximum generations*, the process will end when a specified number of generations have evolved. By the *elapsed time*, the process will end when a specified time has elapsed. According to the *no change in fitness*, the process will end if there is no change in the best fitness of the population for a specified number of generations. Finally, by the *reach the solution or fitness*, the process will end when a previously defined solution or fitness has been reached.

MyGA and Modified MyGA

The original MyGA program uses a population of 10. The niching, scaling, elitism and stochastic roulette are implemented in the selection operator. The crossover operator is the simulated binary crossover (SBX), with $p_c = 0.6$. The mutation operator is the nonuniform, with $p_m = 0.005$. The termination is achieved with 2000 generations.

The modified MyGA (mMyGA) is implemented with the following modifications: initial population greater than current population, niching/tournament selection, baricentric random crossover ($p_c = 0.9$), modified non-uniform mutation ($p_m = 0.1 - 0.4$), and adaptive maximum generations termination.

The fluxogram of mMyGA is shown in Figure 1, and the modifications are fully explained below.

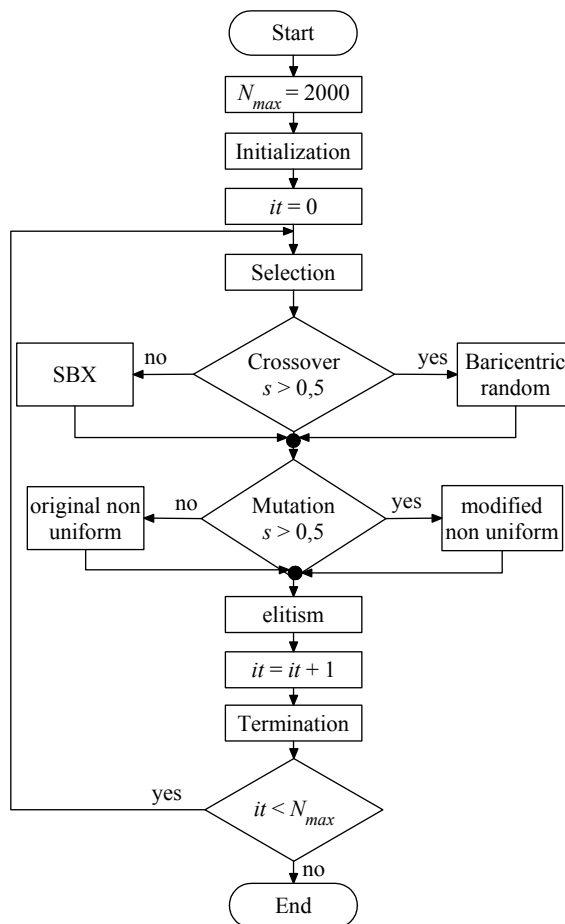


Figure 1: Fluxogram of mMyGA

a) Initial Population

The initial population is 50 times the population (N_{pop}), while N_{pop} is calculated as 10 times the number of variables. The initial chromosomes are created at random and sorted by fitness, and the best N_{pop} chromosomes constitute the population for the starting point. The advantage of this approach is that, by creating a large initial population, one can initially cover a greater amount of the solution search space and then use a population containing relatively strong chromosomes.

b) Selection

The niching selection compares the fitness values of two solutions from 25% of the population within a normalized Euclidian distance (d_{ij}). If $d_{ij} < 0.1$, then a tournament selection is performed. This selection consists in picking two solutions at random and comparing their fitness values, and the winner is duplicated.

c) Crossover

The baricentric random (Yedder, 2002) and SBX (Deb, 2000) methods are used at random in order to improve the crossover. When the crossover operator is selected, a random number between 0 and 1 is generated: if this number is less than 0.5, the SBX method is used, but otherwise, the baricentric random method is used; in this way, there is a greater variety among the offspring. Also, the brood recombination is implemented.

d) Mutation

The non-uniform mutation (Michalewics, 1996; Yedder, 2002) and an empirical modification are used at random. When the mutation operator is selected, a random number s is generated: if $s \leq 0.5$, the original nonuniform is used, but otherwise, the modified nonuniform is used. The original nonuniform uses another random number, q , to choose between two functions:

$$x_i^n = \begin{cases} x_i + \Delta(t, (x_{\max i} - x_i)) & \text{if } q \leq 0.5 \\ x_i - \Delta(t, (x_i - x_{\min i})) & \text{if } q > 0.5 \end{cases} \quad (1)$$

while the modification includes q inside the function:

$$x_i^n = \begin{cases} x_i + \Delta(t, q(x_{\max i} - x_i)) & \text{if } q \leq 0.5 \\ x_i - \Delta(t, q(x_i - x_{\min i})) & \text{if } q > 0.5 \end{cases} \quad (2)$$

The function $\Delta(t, x)$ was originally defined in Yedder (2002). Also, a variable mutation probability p_m is used (between 0.1 and 0.4, with increments of 0.05); according to Srinivas and Patnaik (1994), in this way it is possible to maintain

diversity in the population.

e) Termination

The procedure terminates when the fitness value does not vary for 1000 generations. When $it = N_{\max}$, the fitness value is compared with the value for $it = N_{\max} - 1000$. If the fitness value has improved, N_{\max} is increased by 1000 and the algorithm continues. For each N_{\max} increment, the mutation probability p_m is increased by 0.05; if $p_m > 0.4$, then it returns to the initial value of 0.1. This is known as a no change in fitness.

In summary, Table 1 shows the setting used for the mMyGA program.

Table 1: Parameters used in the mMyGa program

Parameter	mMyGA
Initial population	50 times population
Population	10 times variables
Selection	niching/tournament/elitism
Crossover	baricentric random and SBX
Mutation	nonuniform, original/modified
Termination	no change in fitness

VALIDATION AND RESULTS

Validation

In this section, the mMyGA was validated using the examples from Gau et al. (2000); these authors used interval analysis for the estimation of binary energy interaction parameters of the Wilson (1964) model for the activity coefficient by fitting VLE data for the water + formic acid, tert-butanol + 1-butanol and water + 1,2-ethanediol systems. All experimental data were taken from DECHEMA VLE Data Collection (Gmehling et al., 1981). The objective function (FO) used is given by

$$FO = \sum_{j=1}^N \left(\frac{\gamma_{1j}^{\text{exp}} - \gamma_{1j}^{\text{cal}}}{\gamma_{1j}^{\text{exp}}} \right)^2 + \sum_{j=1}^N \left(\frac{\gamma_{2j}^{\text{exp}} - \gamma_{2j}^{\text{cal}}}{\gamma_{2j}^{\text{exp}}} \right)^2 \quad (3)$$

where γ_{ij}^{cal} and γ_{ij}^{exp} are the activity coefficient calculated from the Wilson model and experimental values, respectively. The original MyGA program was used without alteration. The results for objective function OF with all methods of optimization are shown in Tables 2, 3 and 4 for all systems studied. The results show that, as a rule, the three methods,

interval analysis, MyGA and mMyGA, achieve the same optimal point. However, for the water + formic acid system (Table 3), the mMyGa program yield better results than the interval analysis for the three data sets from Chalov and Aleksandrova. (1957, 1958) and Takagi (1939).

The interval analysis of Gau et al. (2000) uses the Newton method within an initial interval to get local optima and can be computationally expensive compared to stochastic techniques. For example, the convergence time using interval analysis was greater than that of the simulated annealing method, according to Souza et al. (2004), who used the Peng Robinson (1976) equation of state with two parameters. These claims are confirmed by Ortega and Espiau (2003), who recommend using a genetic algorithm when there are more than two parameters to optimize. Otherwise, the genetic algorithm can be used in problems with greater numbers of parameters.

With the estimated parameters by all methods, the average percent deviations for T , y_1 , and y_2 for all systems, calculated according to Valderrama and Alvarez (2005), eqns. (5) and (6), are reported in Table 5. Figure 2 shows the T- y_1 equilibrium diagram for ter-butanol + 1-butanol (Wisniak and Tamir, 1976); the calculation with DECHEMA parameters shows the highest deviations.

Table 2: Minimization for water-1,2-ethanediol

Ref.	DECH	Gau	MyGA	mMyGA
Trimble and Potts (1935)	3.0513	1.0391	1.0391	1.0391

Table 3: Minimization for water + formic acid

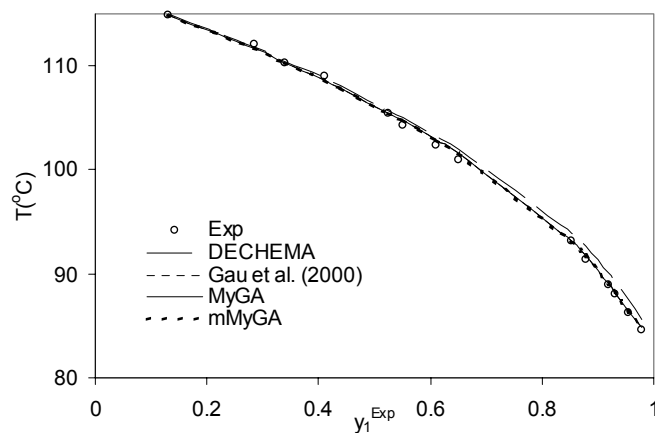
Ref.	DECH	Gau	MyGA	mMyGA
Chalov and Aleksandrova. (1958)	0.0814	0.0814	0.0814	0.0814
	0.1647	0.1114	0.1114	0.1114
	0.1316	0.0858	0.0819	0.0821
	0.0342	0.0342	0.0358	0.0342
Conti et al. (1960)	0.0107	0.0107	0.0358	0.0108
Ito and Yoshida (1963)	0.0399	0.0372	0.0372	0.0372
	0.0459	0.0342	0.0342	0.0342
Melnikov and Tsirlin (1956)	0.0147	0.0109	0.0109	0.0111
	0.3526	0.3526	0.3526	0.3527
Murayama (1961)	0.0257	0.0257	0.0291	0.0258
Plewes et al. (1959)	0.0708	0.0708	0.0708	0.0709
Rivenq (1960)	0.1508	0.0821	0.0822	0.0821
	0.3567	0.3484	0.4235	0.3484
	0.061	0.061	0.061	0.061
	0.0799	0.0799	0.0799	0.0802
Sheinker and Peresleni (1952)	0.1424	0.1424	0.1424	0.1424
Vernon (1939)	0.1606	0.1408	0.1408	0.1409
Chalov and Aleksandrova (1957)	0.1673	0.1153	0.1112	0.1112
	0.1302	0.079	0.2968	0.079
	0.0343	0.0342	0.0358	0.0342
Takagi (1939)	0.0281	0.0282	0.0269	0.0268

Table 4: Minimization for tert-butanol + 1-butanol

Ref.	DECH	Gau	MyGA	mMyGA
Wisniak and Tamir (1976)	0.0333	0.0112	0.0112	0.0112
Zong et al. (1983)	0.1299	0.1164	0.1164	0.1164

Table 5: Mean deviations for T, y₁, and y₂

Method	$ \Delta T $ %	$ \Delta y_1 $ %	$ \Delta y_2 $ %
DECHEMA	0.9	15.5	55.2
Gau et al.	0.8	15.1	41.1
MyGA	0.8	15.0	43.5
mMyGA	0.8	15.1	41.2

**Figure 2:** T-y₁ equilibrium diagram for tert-butanol (1) + 1-butanol (2) at 100 mmHg

It is shown that, even for relatively simple models, such as the Wilson equation, multiple local optima can occur in parameter estimation. It is also shown that, for some data sets, parameter values published in the DECHEMA VLE Data Collection correspond to a local minimum.

$$\%|\Delta T| = \frac{100}{N} \sum_{j=1}^N \left(\frac{|T^{\text{exp}} - T^{\text{cal}}|}{T^{\text{exp}}} \right)_j, \quad (4)$$

$$\%|\Delta P| = \frac{100}{N} \sum_{j=1}^N \left(\frac{|P^{\text{exp}} - P^{\text{cal}}|}{P^{\text{exp}}} \right)_j$$

$$\%|\Delta y_1| = \frac{100}{N} \sum_{j=1}^N \left(\frac{|y_1^{\text{exp}} - y_1^{\text{cal}}|}{y_1^{\text{exp}}} \right)_j, \quad (5)$$

$$\%|\Delta y_2| = \frac{100}{N} \sum_{j=1}^N \left(\frac{|y_2^{\text{exp}} - y_2^{\text{cal}}|}{y_2^{\text{exp}}} \right)_j$$

Applications

a) Vapor-liquid equilibria at high pressure

The maximum likelihood principle (MLE) has been applied by several researchers for the correlation of vapor-liquid equilibrium data under atmospheric pressure up to high pressure, such as Cardozo-Filho et al. (1997). These authors applied MLE to the modeling of the CO₂ + ethanol binary system (Yoon et al., 1993; Suzuki et al., 1990) with the Peng-Robinson equation of state, and their results are better than those with the least squares error (LS) method. Then, the mMyGa program was applied to adjust the same binary system with the simple least squares error in the objective function, as is shown below:

$$FO_{A1} = \sum_{j=1}^N \left(P_j^{\text{exp}} - P_j^{\text{cal}} \right)^2 + \sum_{j=1}^N \left(y_{1j}^{\text{exp}} - y_{1j}^{\text{cal}} \right)^2 \quad (5)$$

where P is the system pressure, y_1 is the molar fraction of the CO₂ in the gas phase, and the

superscripts “*exp*” and “*cal*” are the experimental and calculated values respectively.

The critical properties and acentric factor were taken from Diadem Public 1.2, and the results show what mMyGa can adjust the model better than MLE, as can be seen in Table 6 and Figure 3.

b) Vapor-Liquid Equilibria at High Pressure With Restrictions: Ionic Liquid Containing Systems

The ionic liquid containing CO₂ + 1-n-butyl-3-methylimidazolium hexafluorophosphate ([bmim][PF₆]) system, with VLE data determined by Pérez-Salado Kamps et al. (2003) at 293 K up to 393 K, were correlated by using the Peng-Robinson (1976) equation of state with the Wong-Sandler (1992) mixing rule. The NRTL model (Renon and Prausnitz, 1968) was used to calculate the excess Gibbs free energy in the Wong-Sandler mixing rule. The binary interaction energy and the nonrandomness parameters for the NRTL model (α_{ij} , u_{ij} , u_{ji}) as well as the binary interaction parameter of the Wong-Sandler mixing rule (k_{ij}) were estimated as temperature-independent parameters through the mMyGA method, with a least-squares objective function:

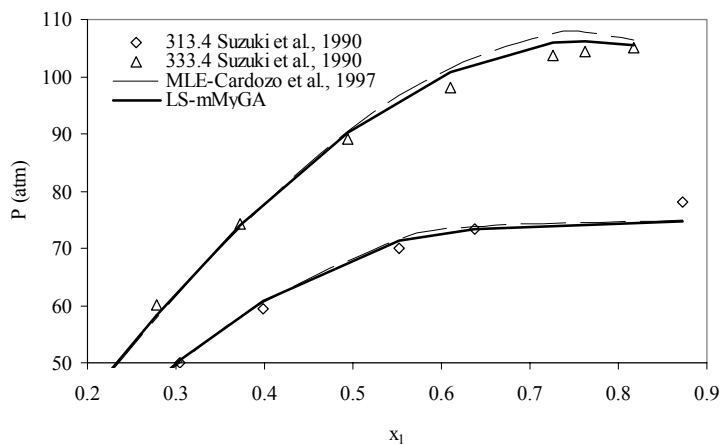
$$FO_{A2} = \sum_{j=1}^N \left(P^{\text{exp}} - P^{\text{cal}} \right)_j^2 + \sum_{j=1}^N \left(1 - y_1^{\text{cal}} \right)_j^2 \quad (6)$$

where P is the system pressure, y_1 is the molar fraction of the CO₂ in the gas phase, and the superscripts “*exp*” and “*cal*” are the experimental and calculated values respectively.

For a good fit of the data set, the concentration of ionic liquid ($1 - y_1^{\text{cal}}$) and the deviations in pressure must be low. The accepted values of $(1 - y_1^{\text{cal}}) < 10^{-3}$ and $\% \Delta P < 10$ are used as constraints for every data point in the minimization method. These restrictions are easily put into the mMyGA program, as restrictions for accepting the new offspring in the crossover and mutation operators. The critical properties, acentric factor, and results are shown in Table 7, where the maximum value of the ionic liquid fraction molar in the gas phase is $8 \cdot 10^{-4}$ with $\% \Delta P = 3.3$. The experimental and calculated P-x₁-y₁ diagram for the binary systems at 333.15 K is shown in Figure 4.

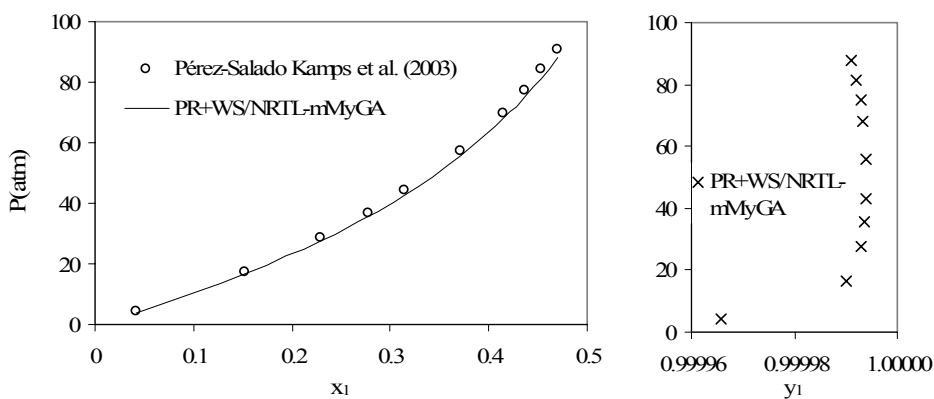
Table 6: Mean deviations for P, y_1 , and y_2 for the CO₂ + ethanol binary system

Method	Interaction parameter	$ \Delta P $ %	$ \Delta y_1 $ %	$ \Delta y_2 $ %
MLE	$k_{ij} = 0.09048, l_{ij} = -0.01414$	4.5	0.4	10.8
mMyGA	$k_{ij} = 0.08844, l_{ij} = -0.01831$	4.6	0.3	10.3

**Figure 3:** Comparison of the different methods: MLE and least squares with mMyGA**Table 7: Critical properties and mean deviations for P and maximum value of ionic liquid in gas phase in the CO₂ + [bmim][PF₆] binary system**

Component	Tc (K)	Pc (atm)	ω
CO ₂ ^a	304.21	72.865	0.2236
[bmim][PF ₆] ^b	708.90	17.074	0.7553

$k_{ij} = 0.5827, (\tau\phi = 0.2585, \nu\tau\phi = 27238.279 \text{ K}\theta/\text{K}\mu\text{o}\lambda, \nu\phi\tau = -1156.827 \text{ K}\theta/\text{K}\mu\text{o}\lambda, \%(\Pi = 3.3, \mu\alpha\xi \psi_2 = 8.10-4$

^aDiadem Public v1.2, ^b Valderrama and Robles (2007)**Figure 4:** P- x_1 - y_1 equilibrium diagram for CO₂ (1) + [bmim][PF₆] (2) at 333.15 K

CONCLUSIONS

Binary vapor-liquid equilibrium data for three systems up to 760 mmHg including polar components were correlated by the Wilson model as a validation of the mMyGA program. The results were compared with the DECHEMA VLE Data Collection and interval analysis. The results indicate that mMyGA is equivalent to interval analysis. It can be concluded that, even for relatively simple models such as the Wilson equation, multiple local optima can occur in parameter estimation. It is also concluded that, for some data sets, parameter values published in the DECHEMA VLE Data Collection correspond to a local minimum.

Two applications of mMyGA have been carried out, one of them to the CO₂ + ethanol system and the other one to the CO₂ + ([bmim][PF₆]) system. In both of cases, the results obtained when the different parameters were estimated with mMyGA were satisfactory.

REFERENCES

- Alvarez, V. H. Modeling of Vapor-Liquid Equilibrium in Mixtures containing Ionic Liquids, M.Sc. Thesis (in Portuguese), School of Chemical Engineering, State University of Campinas: Campinas, Brazil (2007).
- Banzhaf, W., Nordin, P., Keller, R. E., Francone, F. Genetic Programming. An Introduction on the Automatic Evolution of Computer Programs and its Applications, San Francisco, CA: Morgan Kaufmann Publishers, Inc., 1998.
- Cardozo-Filho, L., Stragevitch, L., Wolff, F., Meireles, M. A. The generalized maximum likelihood method applied to high pressure phase equilibrium. *Ciênc. Tecnol. Aliment.*, 17 (1997), 481-484.
- Chalov, N. V., Aleksandrova, O. A. Liquid-vapor phase equilibrium of formic acid-water at normal and low pressures (in Russian), *Gidroliz. i Lesokhim. Prom.*, 10 (1957), 15-17.
- Chalov, N. V., Aleksandrova, O. A. Liquid-vapor phase equilibria in the formic acid-water system (in Russian), *Sbornik Trudov Vsesoyuz. Nauch.-Issledovatel. Inst. Gidroliz. i Sul'fitno-Spirt. Prom.*, 6 (1958), 149-156.
- Conti, J. J., Othmer, D. F., Gilmont, R. Composition of vapors from boiling binary solutions. Systems containing formic acid, acetic acid, water, and chloroform, *J. Chem. Eng. Data*, 5 (1960), 301-307.
- Deb, K. An efficient constraint handling method for genetic algorithms, *Comput. Methods Appl. Mech. Engr.*, 186 (2000), 311-338.
- Diadem Public 1.2. The DIPPR Information and Data Evaluation Manager (2000).
- Gau, C-Y, Brennecke, J. F., Stadtherr, M. A. Reliable nonlinear parameter estimation in VLE modeling, *Fluid Phase Equilib.*, 168 (2000), 1-18.
- Gmehling, J., Onken, U., Arlt, W., Vapor-Liquid Equilibrium Data Collection: Parts 1-8, Frankfurt: DECHEMA Chemistry Data Series (1981).
- Goldberg, D. E. Genetic Algorithms in Search, Optimization and Machine Learning, Reading, MA: Addison-Wesley, 1989.
- Haupt, R. L., Haupt, S. E. Practical Genetic Algorithms, New York: Wiley, 1998.
- Holland, J.H. Adaptation in Natural and Artificial Systems, Ann Arbor, MI: University of Michigan, 1975.
- Hua, J. Z., Brennecke, J.F., Stadtherr, M.A. Enhanced interval analysis for phase stability: Cubic equation of state models, *Ind. Eng. Chem. Res.*, 37 (1998a), 1519-1527.
- Hua, J. Z., Brennecke, J.F., Stadtherr, M.A. Reliable computation of phase stability using interval analysis: Cubic equation of state models, *Comput. Chem. Eng.*, 22 (1998b), 1207-1214.
- Ito, T., Yoshida, F. Vapor-liquid equilibria of water-lower fatty acid systems: Water-formic acid, water acetic acid and water-propionic acid. *J. Chem. Eng. Data*, 8 (1963), 315-320.
- Melnikov, N. P., Tsirlin, Y. A. The vapor-liquid equilibrium in the water-formic acid system (in Russian), *Zh. Fiz. Khim.*, 30 (1956), 2290-2293.
- Michalewicz, Z. Genetic Algorithms + Data Structures = Evolution programs", New York: Springer, 1996.
- Mitchell, M. An introduction to Genetic Algorithms, Cambridge, MA: MIT Press, 1998.
- Murayama, Y. Vapor-liquid equilibria of acetic acid-formic acid-water (in Japanese). *Nippon Kagaku Zasshi*, 82 (1961), 550-554.
- MyGAprogram, <http://cermics.enpc.fr/~benhaj/MyGa/>, downloaded in June 2005.
- Ortega, J., Espiau, F. A new correlation method for vapor-liquid equilibria and excess enthalpies for nonideal solutions using a genetic algorithm. Application to ethanol + an n-alkane mixtures, *Ind. Eng. Chem. Res.*, 42 (2003), 4978-4992.
- Peng, D. Y., Robinson, D. B. A new two-constant equation of state. *Ind. Eng. Chem. Fund.*, 15 (1976), 59-64.
- Pérez-Salado Kamps, A., Tuma, D., Xia, J., Maurer,

- G. Solubility of CO₂ in the ionic liquid [bmim][PF₆], *J. Chem. Eng. Data*, 48 (2003), 746-749.
- Plewes, A. C., Pei, D. C., Code, K. K. Thermodynamic properties of the system formic acid-water, *Can. J. Chem. Eng.*, 37 (1959), 121-125.
- Renon, H., Prausnitz, J.M. Local composition in thermodynamic excess functions for liquid mixtures, *AIChE J.*, 14 (1968), 135-144.
- Rivenq, F. Ebulliometry of solutions of formic acid and water (in French), *Bull. Soc. Chim. France*, (1960), 1505-1507.
- Sheinker, Y. N., Peresleni, E. M. Phase equilibrium of liquid and vapor in some binary systems at reduced pressures (in Russian), *Zh. Fiz. Khim.*, 26 (1952), 1103-1109.
- Souza, A. T., Corazza, M. L., Cardozo-Filho, L., Guirardello, R., Meireles, M.A. Phase equilibrium measurements for the system clove (*Eugenia caryophyllus*) oil + CO₂. *J. Chem. Eng. Data*, 49 (2004), 352-356.
- Srinivas, M., Patnaik, L. M. Adaptive probabilities of crossover and mutation in genetic algorithms, *IEEE Trans. Sys. Man Cyb.*, 24 (1994), 656-667.
- Sun, A. C., Seider, W. D. Homotopy-continuation method for stability analysis in the global minimization of the Gibbs free energy, *Fluid Phase Equilib.*, 103 (1995), 213-249.
- Suzuki, K., Sue, H., Itou, M., Smith, R.L., Inomata, H., Arai, K., Saito, S. Isothermal vapor-equilibrium data for binary systems at high pressure: Carbon dioxide-methanol, carbon dioxide-ethanol, carbon dioxide-1-propanol, methanol-ethanol, methanol-1-propanol, ethane-ethanol, and ethane-1-propanol systems. *J. Chem. Eng. Data.*, 35 (1990), 63-66.
- Takagi, S. Boiling points of the formic acid and water system (in Japanese), *Bull. Chem. Soc. Jpn.*, 14 (1939), 508-509.
- Trimble, H. M., Potts, W. Glycol-water mixtures vapor pressure-boiling point-composition relations, *Ind. Eng. Chem*, 27 (1935), 66-68.
- Valderrama, J. O., Alvarez, V. H. Correct way of reporting results when modelling supercritical phase equilibria using equations of state. *Can. J. Chem. Eng.*, 83 (2005), 578-581.
- Valderrama, J. O., Robles, P. A. Critical properties, normal boiling temperatures, and acentric factors of fifty ionic liquids. *Ind. Eng. Chem. Res.*, 46 (2007), 1338-1344.
- Vernon, A. A. An experiment on liquid-vapor equilibrium for a two-component system, *J. Chem. Educ.*, 16 (1939), 20-21.
- Wilson, G. Vapor-liquid equilibrium. XI. A new expression for the excess Gibbs energy of mixing. *J. Am. Chem. Soc.*, 86 (1964) 127-130.
- Wisniak, J., Tamir, A. Correlation of the boiling points of mixtures, *Chem. Eng. Science*, 31 (1976), 631-635.
- Wong, D. H. S., Sandler, S. I. A theoretically correct mixing rule for cubic equations of state. *AIChE J.*, 38(1992), 671.
- Yedder, A. B. H. Numerical Optimization and Optimal Control: Applications in Molecular Chemistry, Ph.D. diss. (in French), Ecole National des Ponts et Chaussées (2002).
- Yoon, J. H., Lee, H. S., Lee, H. High-pressure vapor-liquid equilibria for carbon dioxide + methanol, carbon dioxide + ethanol, and carbon dioxide + methanol + ethanol, *J. Chem. Eng. Data.*, 38 (1993), 53-55.
- Zong, Z. L., Yang, X. H., Zheng, X. Y. Determination and correlation of vapor-liquid-equilibria of alcohol solutions, *J. Chem. Eng. Japan*, 16 (1983), 1-6.
- Zhu, Y., Xu, Z. Lipschitz optimization for phase stability analysis: Application to Soave-Redlich-Kwong equation of state, *Fluid Phase Equilib*, 162 (1999), 19-29.