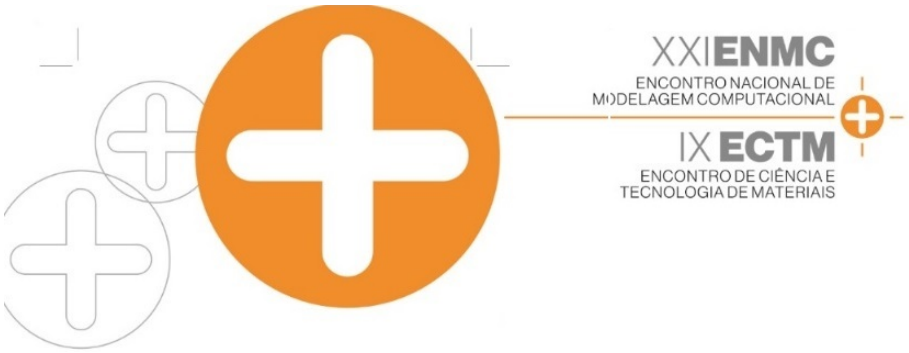

08 a 11 de Outubro de 2018
Instituto Federal Fluminense
Búzios - RJ

Application of Differential Evolution and PA-jDE in chemical engineering problem

Vinícius Magno de Oliveira Coelho[1] - vcoelho@iprj.uerj.br
Francisco Duarte Moura Neto[1] - fmoura@iprj.uerj.br
Gustavo Mendes Platt[2] - gmplatt@furg.br
[1]Universidade do Estado do Rio de Janeiro, Instituto Politécnico - Nova Friburgo, RJ, Brazil
[2]Universidade Federal do Rio Grande - Santo Antônio da Patrulha, RS, Brazil

Abstract. *In recent years, researchers have used optimization techniques to solve problems modelling real-world situations. Among these techniques, the ones that are most suitable for the complexity of these problems are stochastic optimization techniques. The algorithm of the Differential Evolution (DE) has shown itself suitable in its applicability in problems that describe real behaviors; this fact is an incentive for the researchers to look for means to improve its robustness and minimizing computational cost, using methods such as adaptive parameters, adaptive population. Researchers have presented results where it is noticeable the applicability of Differential Evolution in engineering problems. Therefore, this work aims to apply DE and a variant, Self-adapting Differential Evolution algorithm with population Adaptation (PA-jDE), in problems pertaining to the Chemical Engineering field. The results revealed a more rapid convergence of PA-jDE when compared to DE.*

Keywords: *Differential Evolution, Self-adapting Differential Evolution algorithm with population Adaptation,Chemical Engineering problems*

1. Introduction

Mathematical models that describe mundane processes exhibit a high complexity, such as nonlinearity and discontinuity. To address these problems, from the optimization, the most feasible choice belong to the class of stochastic methods. Stochastic methods are characterized by their probabilistic virtue, and it is not necessary to obtain information about the gradient of function, among these methods those that stand out for their results and high applicability are the evolutionary Algorithms. Among the genetic algorithms the Differential Evolution presents satisfactory results, being widely used by the researchers. The Differential evolution (DE) is a parallel direct search method proposed by Storn & Price (1994).

The erroneous adjustment of the parameters of the Differential Evolution algorithm can result in unwanted behaviors, such as increasing the computational time or the premature convergence of the method. In order to refine the results, reduce the number of evaluations of the objective function and consequently the computational cost, researchers have proposed certain

modifications in the Differential Evolution algorithm, creating mechanismics adaptatives for the Differential Evolution.

The Self-Adaptive Differential Evolution (SaDE) (Qin et al., 2009), uses adaptive strategies for the generation of trials vectors and process of adaptation of the CR, F and K parameters during the iterations of the SaDE algorithm. From a previously fixed number called learning period (LP) SaDE uses the probability factor p to choose which mutation strategy to use. This factor is calculated from the success and failure scores obtained by the kth strategy. These modifications use stored information from previous iterations, of this form strategies with greater effectiveness are used during the method search process.

Self-Adaptive Differential Evolution with Neighborhood Search (SaNSDE) (Yang et al., 2008), is formulated from the mechanisms of the NSDE (Yang et al., 2007) and SaDE algorithms. It incorporates the self-adaptation of differents mutations strategies, similarly to that used in SADE. The SaNSDE uses a self-adaptive Cr-weighted crossover rate, this strategy is similar to that used by SaDE when recording the CR values that were successful. The value of fitness improvement is also recorded, thus using these values to obtain the new weighted CR, thus resulting in the auto-adaptation of the CR parameter. The parameter F uses different search ways of different neighborhood search operators without adopting self-adaptive processes.

As shown in the algorithms previously mentioned, modifications in the mutation processes and in the ED parameters present satisfactory results in problems where there are several local results. In this sense, the Self-adapting Differential Evolution algorithm with population Adaptation (PA-jDE) proposed by Yang et al. (2013), uses a scheme to adapt the population, the parameters CR and F and the mutation strategy. The Population adaptation was added to increase population diversity, regardless of size. so if the algorithm gets a certain number of iterations stagnant there is the adaptation of the population around the best individual.

In order to analyze the results statistically, the Mann-Whitney Test was used. In this way, it is possible to evaluate if the PA-jDE presents statistical difference in comparison with the canonical DE. In both algorithms the version of DE used is called DE/rand/1/bin.

2. Self-adapting Differential Evolution algorithm with population Adaptation (PA-jDE)

The algorithm PA-jDE (Yang et al., 2013) emerges from the junction of the Population Adaptation scheme with the algorithm proposed by Brest et al. (2006), the Self-Adapting Differential Evolution Algorithm (jDE). The main objective of the PA-jDE algorithm is to have the capacity to increase the population diversity in case of stagnation, thus increasing its exploratory capacity. This characteristic is very necessary in problems where there are minimum locations and it is desired to obtain the global minimum.

For this, the PA-jDE was proposed based on the algorithm of the Differential Evolution, this algorithm has presented excellent results over the years in the most diverse fields of application. Adjacent to the Differential Evolution is adopted an adaptation in the population, if it remains stagnant, and adaptations in the CR and F parameters of Differential evolution. Due to the process of adaptation of the population a procedure is applied so that all the individuals generated during the execution of the algorithm remain in the domain of the problem addressed.

2.1 Differential Evolution

Differential Evolution (DE) is an algorithm that uses vector differences and comparisons between individuals as the main theoretical framework in their formulation. It has been proposed

by Storn & Price (1994) and has since been widely applied in the field of optimization, with several applications in engineering problems, among which one can cite Wang et al. (2008), Lobato et al. (2008) and Regulwar et al. (2010).

The population evolution proposed by Differential Evolution follows three fundamental steps: mutation, crossover and selection. The Differential Evolution algorithm begins its iterative process generating an initial population with NP individuals, the individuals are randomly distributed within the domain of the problem addressed.

In this work we used the DE/rand/1/bin version, as presented in the paper Storn & Price (1994), in this version we add the weighted difference of two individuals to a third individual, it is important to point out that these individuals are different from each other. The mathematical formulation that expresses this process is as follows:

$$v_j^{(G+1)} = x_{r_1}^{(G)} + F\left(x_{r_2}^{(G)} - x_{r_3}^{(G)}\right) \ , \tag{1}$$

where $j = 1, \ldots, \ NP$. In the rand/1/bin scheme, used as basis in this work, three distinct indices and different from the running index j, has aforementioned, denoted by r_1, r_2 and r_3. This indices are selected randomly, respecting the number of individuals of the population (NP). Then, each individual of the trial vector, resulting from the mutation step, emerges from the operations performed on the individuals of the original population represented by these three selected indices. In addition, F represents the amplification factor, which controls the contribution added by the vector difference. According to Storn & Price (1994), $F \in [0, \ 2]$.

The crossover step is introduced to increase the diversity of the individuals resulting from the mutation process. For this, the parameter CR, denominated as crossover rate, is used. According to Storn & Price (1994), $CR \in [0, \ 1]$. As an initial choice it is indicated that $CR = 0.1$. The crossover process is given as

$$u_{jk}^{(G+1)} = \begin{cases} v_{jk}^{(G+1)}, \text{ if } randb\,(k) \leq CR \text{ or } k = rnbr\,(j) \\ x_{jk}^{(G)}, \text{ if } randb\,(k) > CR \text{ and } k \neq rnbr\,(j) \end{cases} \ , \tag{2}$$

such as $k = 1, \ldots, \ d$, where d denotes the dimension of the problem and $randb\,(k) \in [0, \ 1]$ is a random real number with uniform distribution. The choice of the attributes of a given individual is defined by the crossover coefficient, represented by CR, such that $CR \in [0, \ 1]$ is a constant parameter defined by the user. In turn, the term $rnbr\,(j) \in [1, \ d]$ is a randomly chosen index and, according to Storn & Price (1994), its purpose is to ensure that at least one element of $v_j^{(G+1)}$ must be selected to compose the trial vector $u_j^{(G+1)}$.

It is important to emphasize that if any component of the individual vector is outside the predetermined limit of the problem addressed, then it is necessary to use lateral boundary conditions, so that all individuals belong to the problem domain.

After performing the mutation and crossover procedure the individuals are compared among the trial vectors $u_j^{(G+1)}$ and individuals of the initial population of current generation $x_j^{(G)}$, $j = 1, 2, ..., NP$. The condition for evolution of a given individual is related to the value of the objective function to be optimized. If the value of objetive function of $u_j^{(G+1)}$ is better than $x_j^{(G)}$, then the individue $x_j^{(G+1)}$ is updated as $u_j^{(G+1)}$; otherwise the $x_j^{(G+1)}$ is updated as $u_j^{(G)}$.

2.2 The Population Adaptation

The adaptation of the population is applied when the population is stagnant to a certain integer of generations, denominated UN according to the authors Yang et al. (2013), this one is chosen a priori by the user. In this procedure three parameters are used, being lambda d_G, z_G e $\lambda_{j,G}$.

The parameter d_G is the sum Euclidean distances between individuals of a population at the G-th generation. The value of the d_G does not change when the population converges to a point, this point can be the global optimum or a point of stagnation. The d_G is calculated as follow:

$$d_G = \sum_{i_1=1}^{Np} \sum_{i_2=1}^{i_1-1} \sqrt{\sum_{j=1}^{D} (x_{i1,j,G} - x_{i2,j,G})^2}. \tag{3}$$

The parameter z_G is the parameter of control of when the population adaptation will be executed, this parameter is activated if the population is stagnated UN generations. As a counter of stagnant generations, the parameter $\lambda_{j,G}$ is used, at each stagnant iteration there is an increment in its value, as shown below:

$$z_G = \begin{cases} 1, \text{ if } (\lambda_{j,G} \geqslant UN) \\ 0, \text{ otherwise} \end{cases} \tag{4}$$

where

$$\lambda_{j,G} = \begin{cases} \lambda_{j,G-1}, \text{ if } (d_G = d_{G-1}) \\ 0, \text{ otherwise} \end{cases} \tag{5}$$

In this paper the $UN = NP$ at according of Yang et al. (2013). When $z_g = 1$ a new population is generated.

$$\begin{aligned} x_{i,j,G+1} = low_{j,G} + (up_{j,G} - low_{j,G}) rand(N_{j,G}), \\ j = 1, 2, ..., D. \end{aligned} \tag{6}$$

where

$$low_{j,G} = min(m_{j,G}, x_{low,j}) \tag{7}$$

$$up_{j,G} = max(m_{j,G}, x_{up,j}) \tag{8}$$

$x_{low,j}$ and $x_{up,j}$ are, respectively, the lower and upper boundary of the dimension j; $m_{j,G}$ is the value of best element of dimension j; $rand(N_{j,G})$ is a random number with normal distribution with mean $\mu_{j,G}$ and variance $\sigma^2_{j,G}$. The value of $\mu_{j,G}$ and $\sigma^2_{j,G}$ are as follows:

$$\mu_{j,G} = \frac{m_{j,G} - low_{j,G}}{up_{j,G} - low_{j,G}} \tag{9}$$

$$\sigma_{j,G} = \left(1 - \frac{k}{MaxFEs + 1}\right) \overline{\sigma}_{j,G} \tag{10}$$

where

$$\overline{\sigma}_{j,G} = max(\mu_{j,G}, 1 - \mu_{j,G}) \tag{11}$$

$MaxFEs$ is a predefined maximal value of evaluation of objetive function and k is the number of current evaluation of objetive function.

2.3 Adaptation of CR and F

The adaptation of the parameters F and CR follow the approach adopted in jDE Brest et al. (2006). The best parameters lead to better individuals who are more susceptible to generating a better offspring quality. These adaptations follow the following expressions:

$$F_{i,G+1} = \begin{cases} F_l + rand_1 F_u, \text{ if } (rand_2 < \tau_1) \\ F_l, \text{ otherwise} \end{cases} \tag{12}$$

$$CR_{i,G+1} = \begin{cases} rand_3, \text{ if } (rand_4 < \tau_2) \\ CR_{i,G}, \text{ otherwise} \end{cases} \tag{13}$$

where $rand_j$, $j = 1, 2, 3, 4$ are uniform random numbers. The parameters τ_1 and τ_2 are the parameters of probabilities of adjusting the parameters F and CR, respectively. The parameters $F_l = 0,1$ and $F_u = 0,9$, are respectively, initial and final limits of F.

If a vector is generated that does not belong to the problem search space during the mutation procedure a new vector is generated from the following form Yang et al. (2013):

$$u_{i,j} = \begin{cases} (F_{max} x_{up,j} - x_{low,j} + u_{i,j})/F_{max}, \text{ if } (i_{i,j} < x_{low,j}) \\ (F_{max} x_{low,j} - x_{up,j} + u_{i,j})/F_{max}, \text{ if } (i_{i,j} > x_{up,j}) \end{cases} \tag{14}$$

where F_{max} is the maximum value of F, in this paper we adopted $F_{max} = 1,0$.

3. Problem of Parameter Estimation in Vapor-Liquid Equilibrium

The problem of Parameter Estimation in liquid-vapor equilibrium discussed in this paper is to estimate the parameters of the wilson model for the binary system composed of tert-butanol + 1-butanol using the equilibrium data published by Wisniak & Tamir (1976).

The formulation of the objective function used in this work is the formulation of least squares, , it was used this formulation due to the fact that it is widely used in thermodynamic models for vapor-liquid equilibrium data (Bonilla-Petriciolet et al., 2010) . The objective-function is presented as

$$f(\theta_1, \theta_2) \equiv \sum_{j=1}^{np} \sum_{i=1}^{c} \left(\frac{\gamma_{ji,exp} - \gamma_{ji,calc}(\theta_1, \theta_2)}{\gamma_{ji,exp}} \right)^2, \tag{15}$$

where c is the number of components of the mixture, $\gamma_{ji,exp}$ and $\gamma_{ji,calc}$ are, respectively, the experimental and calculated values for the activity coefficients for the component i at the experimental point j and np represents the number of experimental points.

The activity coefficients for the Wilson Model for a binary system are

$$\begin{aligned} \ln(\gamma_{1,calc}) &= -\ln(x_1 + \Lambda_{12} x_2) + x_2 \left[\frac{\Lambda_{12}}{x_1 + \Lambda_{12} x_2} - \frac{\Lambda_{21}}{\Lambda_{21} x_1 + x_2} \right] \\ \ln(\gamma_{2,calc}) &= -\ln(x_2 + \Lambda_{21} x_1) - x_1 \left[\frac{\Lambda_{12}}{x_1 + \Lambda_{12} x_2} - \frac{\Lambda_{21}}{\Lambda_{21} x_1 + x_2} \right], \end{aligned} \tag{16}$$

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

$$\begin{aligned} \Lambda_{12} &= \frac{v_2}{v_1} \exp\left(\frac{-\theta_1}{RT}\right), \\ \Lambda_{21} &= \frac{v_1}{v_2} \exp\left(\frac{-\theta_2}{RT}\right), \end{aligned} \tag{17}$$

where v_1 e v_2 are the molar volumes of the pure components ($v_1 = 94,88$ cm^3/mol and $v_2 = 91,97$ cm^3/mol), θ_1 and θ_2 are the binary parameters to be estimated, R is the universal gas constant (in this paper $R = 1,987$ cal/gmol.K) and T is the absolut temperature in Kelvin.

In order to obtain the values of the experimental activity coefficients it is necessary to use the modified Raoult's Law, this necessity emerges from the addition of the non-ideality of the liquid phase. The modified Raoult's Law differs from Raoult's Law by adding an activity coefficient to describe the non-idealities of the liquid phase. Thus, the equation of the modified Raoult's law is given as:

$$P y_i = x_i \gamma_i P_i^{sat}, \tag{18}$$

where x_i and y_i are the molar fractions of component i in the liquid and vapor phases, respectively, and P_i^{sat} is the vapor pressure of pure component i.

The vapor pressure of pure component i is calculated by Antoine Equation:

$$P_i^{sat} = 10^{A_i - \frac{B_i}{T + C_i}}, \tag{19}$$

where the A_i, B_i e C_i are the parameters of pure components i and T is the temperature in degree Celsius.

From the Eq. (18), using the equilibrium condition, we obtain the expression for the experimental values of the activity coefficients:

$$\gamma_{i,exp} = \frac{y_i P}{x_i P_i^{sat}}, \quad i = 1, 2. \tag{20}$$

This problem in question presents a region with multiple points with small values of objective function, as it is reported by Platt (2016). Among these points are two solutions belonging to this problem, being a global solution and a local solution.

4. Results

In this section we applied the PA-jDE and the canonical DE in its DE/rand/1/bin version in the problem of parameter estimation of the Wilson model for a binary system. Both methods were applied with 100 sequential runs. the stop criteria used in the study were the maximum number of evaluations of the objective function ($10000 \times D$, as recommended in Yang et al. (2013)), or if the distance between the best individual in the population and the overall solution of the problem was lower than a value chosen a priori , adopted as 0.1.

Both methods used the same values for the parameters CR, F and number of individuals in the population (NP), these parameters were chosen as recommended in Storn & Price (1994). In the DE we used the following parameters: $F = 0.5, CR = 0.6, NP = 20$.

The parameters CR, F and NP of PA-jDE have values similar to those used in DE. In this way, the following configuration is presented for the remaining parameters: $F_{max} = 0.9,\ \tau_1 = 0.5,\ \tau_2 = 0.5,\ UN = NP \div 2$. The values of the parameters τ_1, τ_2 were chosen in order to offer a greater probability of performing the adaptation in the parameters CR and F.

The Fig. 1 shows the points obtained by both algorithms with 100 sequential runs. It is important to note that this problem has a global solution and a local solution. As a result of the 100 executions, DE found 69 global solutions and 31 casual solutions, among which are local solutions and outliers. In contrast to the result of DE, PA-jDE obtained 89 global solutions and 11 casual solutions, among which are the local solutions and outliers.

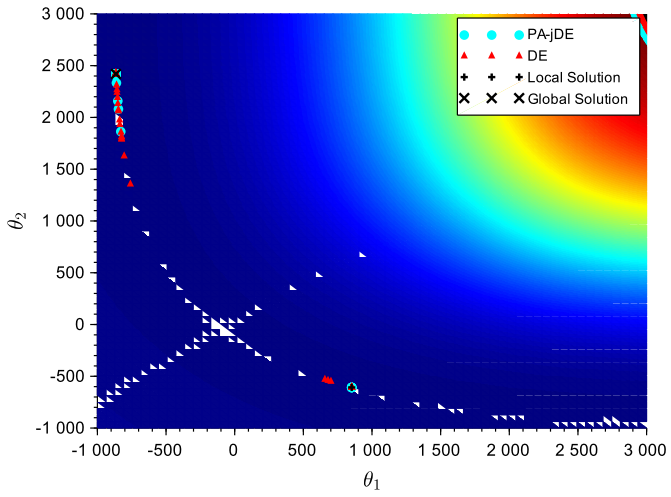

Figure 1- Solutions obtained from the DE and PA-jDE algorithms.

As previously mentioned in this work, the PA-jDE presented better results in relation to the computational cost compared to the DE. The Table 1 presents the average of the number of evaluations of the objective function, the average of iterations (generations) and the average of execution time of both methods. Thus, it was concluded that the PA-jDE presented a faster convergence and with greater effectiveness.

It is important to note that in both methods outliers solutions were obtained, in most cases, when the criterion for stopping the maximum number of evaluations of the objective function (fitness value) was reached. The relatively low average of both methods takes into account the criterion of stopping the distance between the best points and the solution existing in the literature of the problem addressed here. The Fig. 2 shows the number of fitness value evaluations in

Table 1- Average number of function evaluation, iterations and execution time obtained in the 100 runs

	Function Evaluation	Iterations	Execution Time (s)
DE	9230	151	6.3
PA-jDE	7985	98	5.6

Table 2- Results obtained from the 100 runs for both algorithms

	Mean	Standard Deviation	Best	Worst
DE - f	0.0146444	0.0077877	0.011147	0.034596
PA-jDE - f	0.0125131	0.0052935	0.011147	0.033339

each execution of both methods. Note that the DE presents greater peaks of fitness values, this illustrates the obtaining of an outlier solution, in this case the DE. However, it is important to state that both methods had this behavior, however, with different frequencies.

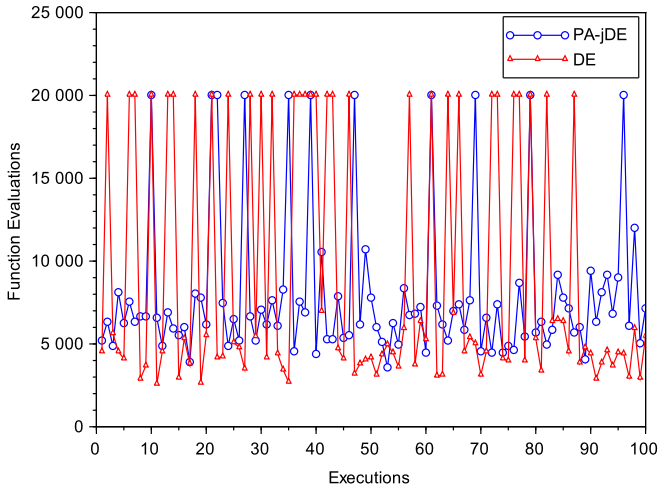

Figure 2- Number of function evaluations at each execution, 100 executions.

The Table 2 shows the results obtained from the 100 runs of the DE and PA-jDE algorithm, it is noticed that the mean and the standard deviation of PA-jDE are lower than the presented by DE, this fact was directly influenced by the number of global solutions found by PA-jDE.

The Fig. 3 presents the results of the objective function (fitness) obtained by both methods at the end of each execution. From this figure it can be analyzed that the DE obtained more solutions outliers in comparison to PA-jDE. This analysis is evidenced by comparing the mean of the objective functions obtained by each method presented in the Table 2.

To evaluate whether there is a significant difference between the results obtained by PA-jde and ED, the Mann-Whitney Test was used. The Mann-Whitney Test is a non-parametric hypothesis test that evaluate whether there is a significant difference between the median of the samples. Due to the fact that the samples are relatively numerous, each sample has 100

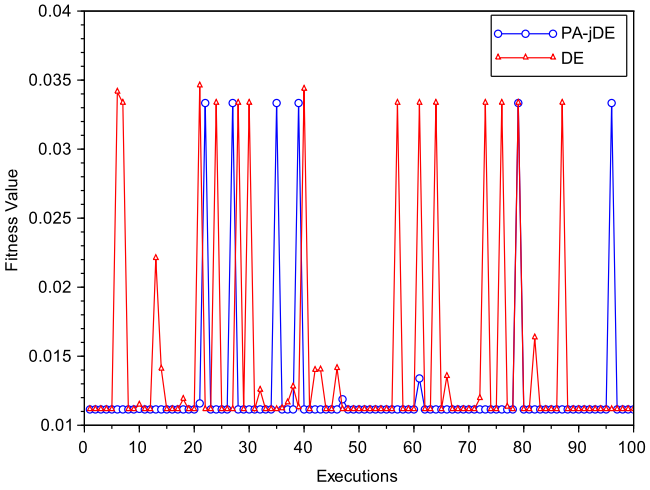

Figure 3- Fitness Value at each execution, 100 runs.

elements, it was necessary to apply the normal approximation and the method was used as two-tails.

From the results the $p-$value obtained is 0.0005 and the value of α is 0.05, where α is a significance level and the value of 0.05 is a standard value of this parameter. Thus, we reject the null hypothesis that states that there is no statistical difference between the medians of the PA-jDE and DE methods.

5. Conclusions

In this work we present the results of the application of the PA-jDE method in a chemical engineering problem comparing the results obtained by the DE. In order to analyze whether there is statistical difference between the medians of these methods, the Mann-Whitney Test was used with α equal 0.05. As a conclusion, for this problem and with the settings presented here, the PA-jDE method has a statistical difference compared to DE and proven by Mann-Whitney Test.

The results obtained by PA-jDE indicate that the application of this method is useful for this type of problem. Although we used a relatively low number of objective function evaluation numbers for a real-world problem, the results were satisfactory.

Consequently, the methods applied in this work presented a good level of applicability in the problem proposed here, being shown as promising methods for solving problems that describe the complexity of the real world.

Acknowledgements

The authors acknowledge the financial support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), from Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) and from Conselho Nacional do Desenvolvimento Científico e Tecnológico (CNPq).

REFERENCES

Bonilla-Petriciolet, A.; Segovia-Hernandez, J. G.; Soto-Bernal, J. J. (2010), "Harmony search for parameter estimation in vapor-liquid equilibrium modeling", in *Proceedings of 5th International Symposium on Design, Operation and Control of Chemical Processes*, 719-726.

Brest, J.; Greiner, S.; Boskovic, B.; Mernik, M.; Zumer, V. (2006), "Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems". *IEEE Transactions on Evolutionary Computation*. vol 10, 646-657.

Lobato, F. S.; Steffen Jr, V.; Arruda, E. B.; Barrozo, M. A. S. (2008), "Estimation of drying parameters in rotary dryers using differential evolution". Journal of Physics: Conference Series, 135, 12063.

Platt, G.M.(2016), Numerical experiments with new metaheuristic algorithms in phase equilibrium problems. International Journal of Mathematical Modelling and Numerical Optimisation. 7, 189-211.

Qin, A.K.; Huang, V.L.; Suganthan, P.N. (2009), "Differential evolution algorithm with strategy adaptation for global numerical optimization", *IEEE Transactions on Evolutionary Computation*, vol 13, 398-417.

Regulwar, D.G.; Choudhari, S.A.; Raj, P.A. (2010), "Differential Evolution Algorithm with Application to Optimal Operation of Multipurpose Reservoir". Journal of Water Resource and Protection. 2, 560–568.

Storn, R.; Price, K. (1997), Differential Evolution – a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization, vol 11, 341-359.

Wang, Z.; Chung, C.Y.; Wong, K.P.; Tse, C. T. (2008), "Robust power system stabiliser design under multi-operating conditions using differential evolution", in *IET Generation, Transmission & Distribution*, vol. 2, 690-700.

Wisniak, J.; Tamir, A. (1976), "Correlation of the boiling points of mixtures". *Chemical Engineering Science*, vol. 31, 631–635.

Yang, M.; Cai, Z.; Li, C.; Guan, J. (2013), "An improved adaptive differential evolution algorithm with population adaptation", in *Proceeding of the fifteenth annual conference on Genetic and evolutionary computation conference - GECCO '13.*, 145-152, ACM Press.

Yang, Z.; Yao, X.; He, J. (2007), "Making a difference to differential evolution". in *Advances in metaheuristics for hard optimization*, Springer, Heidelberg.

Yang, Z.; Tang, K.; Yao, X. (2008), "Self-adaptive differential evolution with neighborhood search", *IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, 1110-1116.