New Strategies Of Differential Evolution For Optimization Of Extraction Process

B.V. Babu* and Rakesh Angira

Department of Chemical Engineering
Birla Institute of Technology and Science
Pilani - 333031 (Rajasthan) India

Abstract

In recent years, evolutionary algorithms (EAs) have been applied to the solution of non-convex problems in many engineering applications. EAs differ from the conventional algorithms since, in general, only the information regarding the objective function is required. In the present work, a test problem on 'optimization of extraction process' is solved using Differential Evolution (DE) and two new DE strategies. The objective of the present study is to maximize the total extraction rate at constant disk rotation speed subject to the inequality constraints. In 1980, scientist used a modified gradient-projection technique and in 1989 GRG (generalized reduced gradient method) was used to solve this problem. Apart from the well known seventh strategy (DE-7) i.e., DE/rand/1/bin, the two new strategies (NS-1 & NS-2) have been applied. A comparison of DE-7 with the proposed two new DE strategies is presented. Experimental Simulations are carried out by running the code for all possible combination of the DE key parameters F & CR (0.0<F<=1; and 0.0<CR<=1.0), where F is scaling factor & CR is crossover constant. The proposed two new DE strategies found to be better and robust ensuring 100% convergence to the global optimum.

Introduction

Liquid-Liquid extraction is a mass transfer operation in which a liquid solution (the feed) is contacted with an immiscible or nearly immiscible liquid (solvent) that exhibits preferential affinity or selectivity towards one or more of the components in the feed. Two streams result from this contact: the extract, which is the solvent rich solution containing the desired extracted solute, and the raffinate, the residual feed solution containing little solute. Liquid-liquid extraction (LLE) is a powerful separation technique and in situations where distillation is not feasible for reasons such as a complex process sequence, high investment or operating costs, heat sensitive materials, or low volatility, extraction is often the best technology to use.

LLE is carried out either (1) in a series of well-mixed vessels or stages (well mixed tanks or plate columns or (2) in a continuous process such as spray columns, packed columns, and rotating disk columns. This example illustrate the application of Evolutionary Computation method such as Differential Evolution (DE) to the optimization of a LLE system represented by a plug flow model. Steady state continuous countercurrent liquid extraction can be modeled in a variety of ways, the most common of which are (1) a plug flow model and (2) an axial dispersion model. Jackson and Agnew [1] demonstrated the effectiveness of an on-line model based steady-state optimization scheme in finding and holding the optimum operation conditions of a liquid extraction pilot plant in which acetic acid was extracted from amyl alcohol using water as the solvent. The equipment could be operated either manually or under computer control. Under automatic operation, the computer could maintain the interface level, feed and solvent flow rates, and the stirrer speed to their respective set points. The latter three set points could be changed by the optimization routine. The interface level was controlled via the extract flow. The measured variables, for both control and optimization purposes were the feed and solvent flow rates; the feed, raffinate and extract concentrations; the stirrer speed; the interface position; and the feed temperature. Further details are given in [2].

The process operation was subject to upper and lower limits on feed and solvent phase superficial velocities and the stirrer speed, and to minimum

* Corresponding Author:
Assistant Dean – ESD & Head of Chemical Engineering & Engineering Technology Departments;
Ph: +91-1596-245073 Ext. 205 / 224;
Fax: +91-1596-244183;
E-mail: bvbabu@bits-pilani.ac.in;
Home Page: http://bvbabu.50megs.com;

throughput and flooding constraints. For use in an optimization scheme, a process model is required to be able to predict the steady-state process output for a given set of inputs. These predicted values can then be used to calculate the value of the performance function. To be of use in an on-line scheme, the model must be amenable to solution without excessive computational effort. They examined the accuracy of four models viz., a model based on plug flow of both phases, one based on axial dispersion superimposed on the flow of both phases, and two empirical models, one linear and one nonlinear) for a continuous pilot-scale extraction column in which water was used to extract acetic acid from amyl alcohol. The linear and non-linear models were direct correlations of experimentally obtained mass transfer data.

A statistical test (an F test or similar) on the model predictions of the process output variables provided a reasonable indication as to whether or not the model would be acceptable for use in a model-based optimization scheme. Two empirical models were rejected, as they were not fitting the data. Both plug-flow and axial-diffusion models enabled the correct optimum to be predicted in all cases, the differences in performance when these models were used were negligible, as there was little axial mixing of the phases in the Rotating Disc Contactor. Also, the marginally superior performance of the axial diffusion model was offset by the greater complexity of its solution compared to the plug flow model. Hence, the plug flow model offered the best compromise between predictive ability and complexity because of its greater simplicity. The same is used in the present study too. Once a model is specified, it can be used to determine the maximum extraction rate. A typical column is shown in fig.1. The DE, process model, the objective function, and the constraints, are described below.

Differential Evolution (DE)

Differential Evolution [3] is an improved version of GA [4] for faster optimization. Unlike simple GA that uses binary coding for representing problem parameters, DE uses real coding of floating point numbers. Among the DE's advantages are its simple structure, ease of use, speed and robustness. Price & Storn [3] gave the working principle of DE with single strategy. Later on, they suggested ten different strategies of DE [5]. A strategy that works out to be the best for a given problem may not work well when applied for a different problem. Also, the strategy and key parameters to be adopted for a problem are to be determined by trial & error. The key parameters of control are: NP - the population size, CR - the crossover constant, F - the weight applied to random differential (scaling factor).

The crucial idea behind DE is a scheme for generating trial parameter vectors. Basically, DE adds the weighted difference between two population vectors to a third vector. Price & Storn [5] have given some simple rules for choosing key parameters of DE for any given application. DE has been successfully applied in various fields. The various applications of DE are: digital filter design [6], batch fermentation process [7, 8], estimation of heat transfer parameters in trickle bed reactor [9], optimal design of heat exchangers [10, 11], synthesis & optimization of heat integrated distillation system [12], optimization of an alkylation reaction [13], scenario-integrated optimization of dynamic systems [14], optimization of non-linear functions [15], optimization of thermal cracker operation [16], a differential evolution approach for global optimization of MINLP problems [17], Optimization of Non-Linear Chemical Processes [18], Evolutionary Computation for Global Optimization of Non-Linear Chemical Engineering Processes [19] etc.

Process Model

Assuming that the concentrations are expressed on a solute free mole basis and that the equilibrium relation between Y and X is a straight, i.e., the phases are insoluble. The model is then given as below:

$$\frac{dX}{dZ} - N_{OX}\left(X - Y\right) = 0 \qquad (1)$$

$$\frac{dY}{dZ} - FN_{OX}\left(X - Y\right) = 0 \qquad (2)$$

where,

F	= extraction factor (mv_X / v_Y)
m	= distribution coefficient (m = 1.5)
N_{OX}	= number of transfer units
v_X, v_Y	= superficial velocity in raffinate, extract phase
X	= dimensionless raffinate phase concentration
Y	= dimensionless extract phase concentration
Z	= dimensionless contactor length

Figure-1 shows the extraction column with boundary conditions X_0 and Y_1. A solution for Y_0 in terms of v_x and v_Y can be obtained, given the values for m, N_{OX}, and the length of the column. Hartland and Mecklenburgh [20] list the solution for the plug flow model (and also the axial dispersion model) for a linear equilibrium relationship, in terms of F:

$$Y_0 = \frac{F\left[1-\exp\left\{N_{OX}\left(1-F\right)\right\}\right]}{1-F\exp\left[N_{OX}\left(1-F\right)\right]} \tag{3}$$

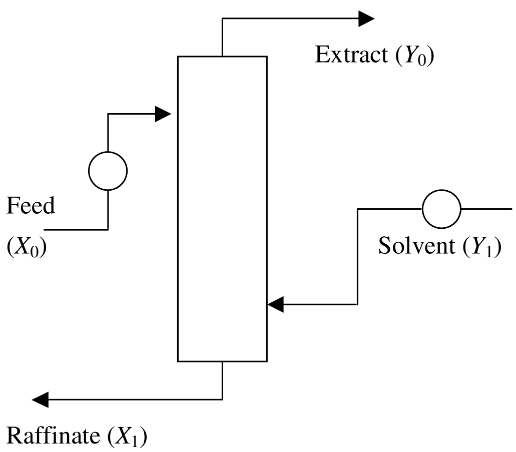

Fig.1: Schematic of Extraction Column

For the plug flow and axial diffusion models, Jackson and Agnew [1] summarized a number of correlations for N_{OX} given by an equation of the form:

$$N_{OX} = a\left(\frac{v_X}{v_Y}\right)^b \left(N\right)^c \tag{4}$$

The correlations obtained by non-linear least-squares regression were:

$$N_{OX} = 1.7\left(\frac{v_X}{v_Y}\right)^{0.24} \left(N\right)^{0.069}$$

for $2.08 \le N \le 4.2s^{-1}$ (5a)

$$N_{OX} = 0.2\left(\frac{v_X}{v_Y}\right)^{0.24} \left(N\right)^{1.5}$$

for $4.2 \le N \le 8.33s^{-1}$ (5b)

$$N_{OX} = 0.18\left(\frac{v_X}{v_Y}\right)^{0.18} \left(N\right)^{0.088}$$

for $2.08 \le N \le 4.3s^{-1}$ (5c)

$$N_{OX} = 0.09\left(\frac{v_X}{v_Y}\right)^{0.24} \left(N\right)^{2.1}$$

for $4.3 \le N \le 8.33s^{-1}$ (5d)

In the present study, equation (5b) is used with $N = 8.33s^{-1}$.

$$N_{OX} = 4.81\left(\frac{v_X}{v_Y}\right)^{0.24} \tag{6}$$

Objective function

We have used the same objective function as proposed in [1] i.e., to maximize the total extraction rate for constant disk rotation speed subject to the inequality constraints:

Maximize $f = v_Y Y_0$ (7)

Inequality constraints

Implicit constraints exist because of the use of dimensionless variables

$X_0 \le X \le X_1$

$Y_1 \le Y \le Y_0$ (8)

Constraints on v_X and v_Y would be upper and lower bounds such as:

$0.05 < v_X < 0.25$

$0.05 < v_Y < 0.30$ (9)

and flooding constraint is:

$v_X + v_Y <= 0.20$

Results & Discussion

Jackson and Agnew [1] used a modified gradient-projection technique for linearly constrained optimization problems developed by [2]. Edgar & Himmelblau [21] used GRG (generalized reduced gradient method) and obtained the same optimum (0.15, 0.05) as Jackson and Agnew. The objective function is 0.225 i.e. the true optimum lay on the flooding constraint.

In the present study, the DE (differential evolution), an evolutionary technique, is applied. Apart from the well known seventh strategy (DE-7) i.e., DE/rand/1/bin, the two new strategies (NS-1 & NS-2) have been applied. They differ slightly from DE-7, the way the mutation and recombination is performed. The psuedo codes for mutation and recombination in *DE-7*, *NS-1 and NS-2* are given below:

(i). For DE-7

```
{
If (rand_j [0,1]) <CR ∨ k = D)
 {
trial[j] = x[c][j]+F*(x[a][j]-x[b][j]);
 }
else trial[j] = x[i][j];
```

```
            j=(j+1)%D;
  /*  % = modulo; index j runs from 0 to D-1 */
            }
(ii). For NS-1
          {
          If (rand_j [0,1]) <CR ∨ k = D)
           {
          trial[j]=x[c][j]+F*(x[a][j]-x[b][j]);
           }
else trial[j]=bestit[j]+F*(x[a][j]-x[b][j]);
            j=(j+1)%D;
          }
(iii). For NS-2:
          {
          If (rand_j [0,1])<CR ∨ k = D)
           {
          trial[j]=x[c][j]+F*(x[a][j]-x[b][j])+(1-
F)*(x[d][j]-x[e][j]);
            }
  else trial[j]=bestit[j]+F*(x[a][j]-x[b][j]);
          j=(j+1)%D;
           }
```

The term *rand*$_j$*[0,1]* represents a uniformly distributed random variable that ranges from zero to one. The subscript *j* indicates that a new random value is generated for each value of *j*; that is, for each object variable. The vector *bestit[j]* is the best vector of the previous generation and the indices *a, b, c, d, e* $\in$ *{1,2,…., NP}* are randomly chosen population indices that are mutually different and also different from *i*. DE employs both mutation and recombination to create one child or trial vector *trial[j]*, for each parent vector *x[i][j]* as shown in psuedo code for DE-7, NS-1 and NS-2.
Figure-2 and figure-3 show the results obtained using DE-7, NS-1 and NS-2 and present the comparison of DE-7 with NS-1 & NS-2. The stopping criterion adopted is to terminate the search process when one of the following conditions is satisfied: (1) the maximum number of generations is reached (assumed 2000 generations). (2) $| f^k_{\max} - f^k_{\min} | < 10^{-5}$ where *f* is the value of objective function for k-th generation. After the mutation & recombination, if the bound (i.e. lower & upper limit of a variable) is violated then it is brought in the bound range (i.e. between lower & upper limit) by randomly assigning a value in the bound range (without forcing).

In Figure-2 & Figure-3, NRC & NFE represent respectively, the mean number of objective function evaluations and the percentage of runs converged to the global optimum in all the 10 executions (with different seed values). The code was run for all possible combination of F & CR (0.0<F<=1; and 0.0<CR<=1.0), where F is scaling factor & CR is crossover constant.

Figure-2 presents NRC vs. CR results obtained for F = 0.5 for all three strategies i.e., DE-7, NS-1 and NS-2. The NRC is more in case of new strategy (NS-2) as compared to DE-7 & NS-1. The new strategy NS-2 seems to be robust than DE-7 & NS-1 due to nearly 100% conversion to global optimum. Also, it is clear that till CR=0.3 the NRC are nil for NS-1 while it is almost 100% for DE-7 & Ns-2. From fig.3, it is evident that NFE are least in case of NS-1. Also, difference of NFE between NS-1 & DE-7 becomes less after CR = 0.5 and both NS-1& DE-7 have same NFE at CR=1.0. NFE are 3.55%, 6.56%, 1.24% more in case of DE-7 than the NS-1 at CR = 0.6, 0.7, 0.9 respectively.

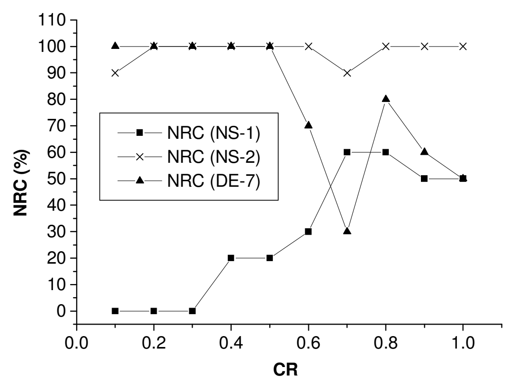

Fig. 2: NRC vs. CR (for F = 0.5)

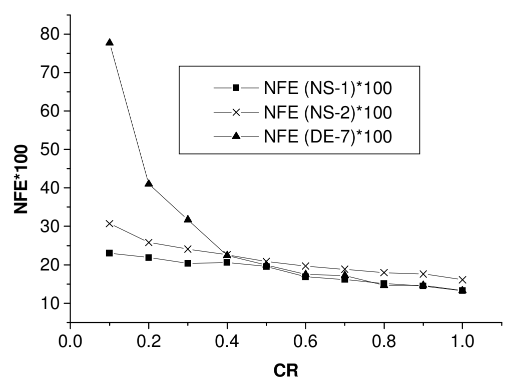

Fig.3: NFE vs. CR (for F = 0.5)

From above figure-3 we can observe that till CR = 0.4, the NFE for NS-2 are less than that of DE-7. But for CR >0.4 the NFE are slightly more than either DE-7 or NS-1. At CR =0.5, the NFE are 6.65% and 4.41% more than that of NS-1 and DE-7 respectively. Although the NFE are slightly more for NS-2 but NRC is nearly 100% indicating the robustness of the strategy.

Conclusions

In the present study, the problem of optimization of liquid extraction process has been solved using DE and two new DE strategies. It is found that the performance of DE-7 is slightly better than that of NS-1. But NS-2 seems to be the best choice and certainly better than DE-7 due to almost 100% NRC which indicate the robustness of this strategy.

References

1. Jackson, P. J. and Agnew, J. B., *Computers & Chemical Engineering, 4*, 241 (1980).
2. Jackson, P. J., Ph.D. Thesis, Monish University, Victoria, Australia (1977).
3. Price, K., and Storn, R., *Dr. Dobb's Journal,* 22 (4)*,* 18 – 24 and 78 (1997).
4. Goldberg, D. E., *Genetic Algorithms in search, Optimization, and Machine learning,* Reading, MA, Addison-Wesley (1989).
5. Price, K. and Storn, R., *Web site of DE as on July, 2003,* the URL of which is: http://www.ICSI.Berkeley.edu/~storn/code.html (2003).
6. Storn, R., *International Computer Science Institute,* TR-95-026 (1995).
7. Chiou, J. P. and Wang, F. S., *Computers & Chemical Engineering,* **23**, 1277-1291 (1999).
8. Wang, F. S., and Cheng, W. M., *Biotechnology Progress,* **15** (5), 949-952 (1999).
9. Babu, B. V. and Sastry, K. K. N., *Computers & Chemical Engineering,* **23**, 327–339 (1999).
10. Babu, B. V. and Munawar, S. A., *Proceedings of All-India seminar on Chemical Engineering Progress on Resource Development: A Vision 2010 and Beyond*, IE (I), Bhuvaneshwar, India, March 11, (2000).
11. Babu, B. V. and Munawar, S. A., *PreJournal.com - The Faculty Lounge, Article No. 003873, posted on website Journal http://www.prejournal.com*. (2001).
12. Babu, B. V. and Singh, R. P., *Proceedings of All-India seminar on Chemical Engineering Progress on Resource Development: A Vision 2010 and Beyond*, IE (I), Bhuvaneshwar, India, March 11, 2000.
13. Babu, B. V. and Aura, C., *Proceedings of International Symposium & 53rd Annual Session of IIChE* (*CHEMCON-2000*), Calcutta, December 18-21 (2000).
14. Babu, B. V. and Gautam, K., *Proceedings of International Symposium & 54th Annual Session of IIChE* (*CHEMCON-2001*), Chennai, December 19-22 (2001).
15. Babu, B. V. and Angira, R., *Proceedings of 12th ISME Conference*, India, January 10–12, 153-157 (2001).
16. Babu, B. V., and Angira, R., *Proceedings of International Symposium & 54th Annual Session of IIChE* (*CHEMCON-2001*), Chennai, December 19-22 (2001).
17. Babu, B. V., and Angira, R., *Proceedings of 4th Asia Pacific Conference on Simulated Evolution and Learning (SEAL-2002)*, Singapore, November 18-22, Vol. 2, pp. 880-884 (2003).
18. Babu, B. V., and Angira, R., *Proceedings of International Symposium & 55th Annual Session of IIChE (CHEMCON-2002)*, OU, Hyderabad, December 19-22 (2002).
19. Angira, R. and Babu, B. V., *Proceedings of International Symposium on Process Systems Engineering and Control (ISPSEC' 03) - For Productivity Enhancement through Design and Optimization,* IIT-Bombay, Mumbai, Paper No. FMA2, pp 87-91 (2003).
20. Hartland, S. and Mecklenburgh, J. C., *The Theory of Backmixing*, Wiley, New York (1975).
21. Edgar, T. F. and Himmelblau, D. M., *Optimization of Chemical Processes,* McGraw-Hill, Inc., Singapore, Page Nos. 534-539 (1989).

Note:

1. References [9, 10, 11, 12, 13, 14, 15, 16, 17, 18 and 19] are also available via Internet as .pdf files at: http://bvbabu.50megs.com/custom.html/#24, 28, 35, 29, 31, 39, 34, 38, 46, 54, 56 respectively.
2. And, references [9, 11, and 14] are also available at Homepage of DE at: http://www.icsi.berkeley.edu/~storn/code.html as Application No. 13, 18, and 21 respectively.