

Standing wave optimization of SMB using a hybrid simulated annealing and genetic algorithm (SAGA)

Fattaneh G. Cauley · Stephen F. Cauley ·
Nien-Hwa Linda Wang

Received: 30 April 2007 / Revised: 18 January 2008 / Accepted: 20 March 2008 / Published online: 8 April 2008
© Springer Science+Business Media, LLC 2008

Abstract In this paper we draw on two stochastic optimization techniques, Simulated Annealing and Genetic Algorithm (SAGA), to create a hybrid to determine the optimal design of nonlinear Simulated Moving Bed (SMB) systems. A mathematical programming model based on the Standing Wave Design (SWD) offers a significant advantage in optimizing SMB systems. SAGA builds upon the strength of SA and GA to optimize the 16 variables of the mixed-integer nonlinear programming model for single- and multi-objective optimizations. The SAGA procedure is shown to be robust with computational time in minutes for single-objective optimization and in a few hours for a multi-objective optimization, which is comprised of more than one hundred points.

Keywords Simulated moving bed · Standing wave design · Simulated annealing · Genetic algorithm · Multi-objective

1 Introduction

Simulated Moving Bed (SMB) is a flexible and powerful separation tool that has been used for the separation of petrochemicals since early 1970s and of sugars since the 1980s

(Broughton 1961; Ruthven and Ching 1989). The application of SMB has been extended into recovery and purification of bio-chemicals and pharmaceuticals (Schulte and Strube 2001; Xie et al. 2001). In an ideal SMB system, where there are no mass-transfer effects, local equilibrium analysis guarantees complete separation, which is achieving 100 percent purity and yield (Mazzotti et al. 1997). However in nonideal systems, where mass-transfer effects are prevalent, high purity and yield in separation are not easily achieved.

Ma and Wang (1997) proposed the Standing Wave Design (SWD) for determining the optimal values of operating variables of nonideal SMB systems with linear isotherms. Here optimal design refers to the determination of the values of five variables (the four zone flow rates and the average velocity of port movement) that maximizes throughput or minimizes the desorbent consumption. Later the SWD was modified for nonideal nonlinear SMB systems (Mallmann et al. 1998; Xie et al. 2003b; Lee et al. 2005b). Furthermore, unlike methods that rely on the equilibrium theory, SWD does not require time-consuming simulations of SMB processes to ensure that the purity and yield requirements are satisfied. Many experimental studies and computer simulations based on a detailed rate model (VERSE) have verified that SWD guarantees the desired purity and yield for nonideal SMB systems with two or more columns in each zone (Wu et al. 1998; Xie et al. 2002, 2003a).

Recent research has considered optimizing additional variables that are thought to have a large impact on the economic efficiency of an SMB unit. Procedures based on SWD (Lee et al. 2005b; Cauley et al. 2004, 2006), and other procedures including the triangle theory (Ludemann-Hombourger et al. 2000; Silva et al. 2004), have been used to determine the optimal particle size, total number of columns, column

F.G. Cauley (✉)
Seaver College, Pepperdine University, Malibu, CA 90263, USA
e-mail: fattaneh.cauley@pepperdine.edu

S.F. Cauley
School of Electrical Engineering, Purdue University,
West Lafayette, IN 47907, USA

N.-H.L. Wang
School of Chemical Engineering, Purdue University,
West Lafayette, IN 47907, USA

lengths, and the feed concentrations using different objectives such as maximum throughput, product purity, yield, and as minimum desorbent consumption.

A notable feature of the SWD is a system of algebraic equations that link product purity and yield to zone lengths, zone interstitial velocities, port-switching time, isotherms, and mass-transfer parameters (axial dispersion and lumped mass-transfer coefficients). Solving this system of equations provides a systematic procedure for achieving the desired purity and yield in the presence of mass-transfer effects and a pressure limit (Ma and Wang 1997; Lee et al. 2005a).

The development of SWD has paved the way for the formulation of a mathematical programming model of the SMB design problem. A formal mathematical programming model of a problem is the initial step toward developing the ability in using standard optimization techniques to solve the problem. Cauley et al. formulated a mixed-integer mathematical programming model of SMB systems, with SWD equations as the core constraints. The problem was then solved by the algorithm SWAT, which is based on a stochastic optimization technique Simulated Annealing (SA), and branch-and-bound method (Cauley et al. 2004, 2006).

The SMB optimization results determined by SWAT were compared to published literature, and showed either similar or improved values. In addition, the SWAT results were obtained in an order of magnitude faster than procedures such as grid search (Lee et al. 2005b; Mun et al. 2003). However, a potential source of error (lack of robustness) in SWAT was the use of a deterministic branch-and-bound algorithm to solve for the SMB configuration. The branch-and-bound method used in the SWAT relies on the solution of “relaxed” optimization problems (non-integer column configurations) in order to “prune” (ignore) regions of the complete search space that are deemed to be irrelevant. This style of algorithm works extremely well for Integer Linear Programming problems. The SMB optimization problem, however, is a nonconvex optimization problem. In this case, the possibility of incorrectly pruning regions is potentially a serious problem that is often difficult to avoid. In our experience, the possibility of incorrectly pruning a region is likely to become more of a problem as the complexity of the constrained optimization increases.

In this paper, a hybrid algorithm SAGA, based on Simulated Annealing and Genetic Algorithm (GA), is introduced. By replacing the deterministic branch and bound in SWAT with the stochastic GA for determining the integer variables of the model, we are presenting an algorithm that is both robust and computationally efficient. SAGA utilizes the best of SA and GA for optimizing the design of SMB systems. As such it is qualitatively different from SWAT or any other stochastic optimization technique.

The results of optimization by SAGA are presented and discussed for both single- and multi-objective optimizations

using an example from the literature. It will be shown that in each optimization the SAGA algorithm determines an SMB system that shows improvement compared to the one determined by SWAT and presented in the previous literature (1–9%). The comparisons lead to the conclusion that the SAGA algorithm offers additional robustness, with the computational time in minutes for single-objective optimization and in a few hours for multi-objective optimization, which is comprised of more than one hundred points. Implementation of SAGA was coded in MATLAB 7.0 and run on a 2.2 GHz Pentium 4.

2 SWD mathematical programming model

The SWD model contains 16 decision variables that are clustered into four tiers to facilitate investigation of different factors that determine the optimal design of an SMB system. Previous papers on the design of SMB systems can be considered as special cases of the SWD model, with the decision variables in some of the tiers regarded as fixed (Xie et al. 2003b; Ludemann-Hombourger et al. 2000). The Tier0 consists of two variables, the feed concentrations $C_{F,1}$, $C_{F,2}$. The two are assumed to be real-valued and distinct. The Tier1 variables determine the construction of the SMB unit and consist of: column length L_c , column cross sectional area S , number of columns N_{col} , and particle diameter d_p . Tier2 variables are made of number of columns N_{col}^j in zone j , for $j = \text{I, II, III}$. The seven Tier3 variables include the five variables that are generally considered as operating variables: interstitial velocity u_0^j in the four zones, and the average velocity of port movement v . The additional two variables, which are unique to nonlinear systems, are the diluted plateau concentrations near the feed port $C_{s,i}$, $i = 1, 2$ (Xie et al. 2003b).

In the SWD model, different SMB systems are represented by different input parameters. The model contains in the set of constraints upper and lower bounds for all variables. The bounds are determined by physical restrictions, maximum concentration limits, and solubility constraints. It should be noted that the Standing Wave Theory is based on true moving-bed processes. For this reason, the standing wave design can not guarantee the desired purity and yield, if each of the four zones has only one column (which in some situations could prove to be the optimal system). However, most of the existing SMB systems have two or more columns in each zone. For such systems, the standing wave design can guarantee the desired purity and yield (Ma and Wang 1997). In order to ensure the validity of the solutions of the above model, a lower bound of two columns is assumed for each zone.

In addition to the SWD equations, the model contains several constraints that represent the requirements of mass

balance at the ports, maximum operating pressure of the system, and the desired annual production. A comprehensive discussion of the other elements of the SWD model, including precise descriptions of the input parameters, decision variables, constraints, and objectives that are considered for optimization, is presented in [Appendix](#).

The SWD model contains a number of constraints that can be described as nonlinear, nonconvex functions of the decision variables. Consequently this model is a mixed-integer nonlinear programming problem and highly nonconvex (Biegler and Grossmann 2004; Luenberger 2003). Thus, it presents a substantial computational challenge.

3 Optimization techniques

Although the field of nonlinear programming has evolved for the last 50 years, no single-solution technique has been recognized as having the capability of providing satisfactory solutions to all problems.

Classical algorithms, such as steepest descent, Newton-Raphson, and quasi-Newton methods, are iterative and are designed to efficiently solve models that satisfy convexity requirements (which our model does not). Classical methods often fail in solving nonconvex problems and will likely stop at local rather than global optima (Luenberger 2003).

Methods that sample the feasible set, both deterministically (e.g., grid search) and stochastically, have been developed as a solution to the problem of obtaining a local rather than a global optima. In these techniques, the feasible set is sampled and the corresponding values of the objective function are calculated. Stochastic optimization techniques, where sampling for the global optima follows a probability distribution, have been developed and are being used with increasing frequency. At the heart of any stochastic optimization algorithm is the opportunity for a jump to another part of the feasible set that is independent of the current point (i.e., acceptable moves do not always have to be associated with better values of the objective function). Among these algorithms, two stand out: Genetic Algorithm (GA) and Simulated Annealing (SA). The former works on the principle of mimicking a biological system, and the latter a physical system.

GA, initially developed by Holland (1975), provides a method, based on survival of the fittest, to intelligently search the feasible set for global optima. The algorithm works loosely as follows: A population (set of feasible points), whose members are called *chromosomes*, is generated randomly. Chromosomes that yield greater fitness (i.e., better values for the objective function) have a better chance of surviving to the next generation. A new generation of chromosomes is generated, either by crossover (splicing together two chromosomes), or mutation (inducing radical

change in a chromosome). Mutations have the ability to take the algorithm to a completely different location in the feasible set, thus escaping local optima. The above steps are repeated until the algorithm is halted. Common halting criteria include stopping when there has been no improvement observed in the value of the objective after many consecutive generations, or after a specified computation time period elapses.

In practice, GA has a strong record of finding a good solution for unconstrained and integer valued problems. However, the efficiency of GA drops substantially in the presence of a large number of constraints (Davis 1987, 1991). The problem arises from the fact that it is possible, and often the case, that the recombination for any two feasible members of a generation is not feasible. There are modifications to GA that transform the constrained optimization problem to an unconstrained problem via a penalty function method. However, such modifications increase the difficulty of the search for feasible points and adversely affect the performance of the algorithm (Michalewicz and Fogel 2000).

During the mid 1980s, scientists at IBM (Kirkpatrick et al. 1983) and Cerny (1985) independently suggested the analogy between annealing and the solution of an optimization problem. The SA algorithm mimics the behavior of a physical system that is heated then cooled slowly, such as growing crystals or annealing metals.

SA works by randomly searching the feasible set looking for low (high) values of the objective function. The search is not exhaustive, but is based on principles of statistical mechanics; it follows the Boltzmann probability distribution for reaching different states. This is equivalent to the algorithm providing opportunities for jumps from current location to other parts of the feasible set to continue the search (Salamon et al. 2002); this in turn helps the algorithm to escape entrapment in local optima and eventually reach the global optima (Michalewicz and Fogel 2000; Salamon et al. 2002).

The literature shows that SA algorithms easily accommodate constraints, a major advantage in optimizing the SWD model. Additionally it does not follow a greedy algorithm approach; rather it "... is not biased by kinetic factors, i.e., by how rapidly we were able to locate a certain point," (Salamon et al. 2002). This property will decrease the likelihood of getting trapped in local optima.

Both algorithms SA and GA have been combined with other procedures to overcome some of the shortcomings mentioned above (Davis 1991). In this paper, a hybrid of the two algorithms will be introduced. SAGA utilizes the best of SA and GA for optimizing the design of SMB systems. As such it is qualitatively different from SWAT or any other stochastic optimization technique.

3.1 SAGA

Figure 1 is a detailed flowchart of the SAGA algorithm. In this hybrid algorithm, determination of the integer portion of the optimization problem (i.e., the total number of columns and the number of columns in zones 1–3) is performed by a GA algorithm, which is illustrated in left portion of the flowchart. For ease of implementation, the GA toolbox provided by MATLAB is employed. The GA carries out a stochastic search through the set of all feasible integer variables within their bounds. A potential set of feasible integer variables constitutes the “chromosomes” from which the initial population is formed.

The GA procedure is carried out with an initial population of five, that is, five different sets of feasible values for the integer variables are considered during each stage of the algorithm. The initial population includes a chromosome corresponding to the lower bound values of the variables (8, 2, 2, 2). The other four chromosomes are constructed by adding one more column to the total number of columns and placing the additional column in each of the four possible zone locations. This can be thought of as starting with the most basic configuration and moving a small amount in

each of these possible “directions.” All chromosomes from the population are then evaluated for fitness using a custom-designed SA algorithm, which is detailed in the right portion of the flowchart.

The short execution time of SAGA (and SWAT) is attributed to the custom-designed SA portion that includes a sophisticated procedure called basin-hopping for the move class. This procedure has been successfully applied to continuous variables, and as the name indicates, consists of moves among local minima by the use of a gradient-based optimizer (Wales and Scherega 1999). A detailed description of our implementation of the SA algorithm is provided in Cauley et al. (2004, 2006). The SA inner loop of the algorithm determines the values of the other 12 variables, such that all constraints are satisfied and a stated objective function is optimized. The fitness of each chromosome is then described by the value of the objective function.

The inclusion of our custom SA routine as a stochastic fitness evaluation for a GA is, to the best of our knowledge, unique in the SMB optimization literature. It is important to note that an inherent degree of flexibility associated with SA is built in this type of fitness evaluation. Given that the SA procedure is used in conjunction with the GA-based search

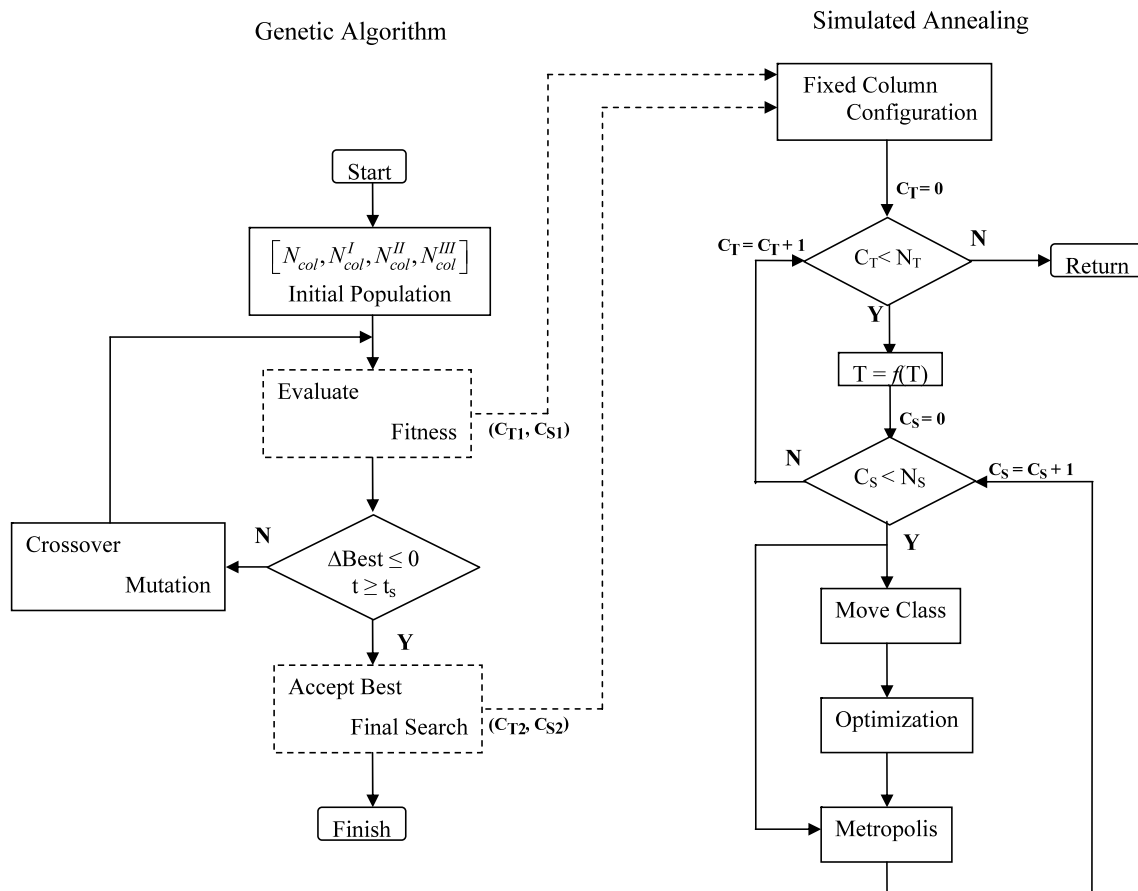


Fig. 1 Flowcharts of the optimization program SAGA

for column configuration, each configuration that exhibits promise may be revisited in future generations.

The basic advantage of SAGA can be seen in the small number of populations that need to be considered in GA. It is important to recall that the GA in SAGA only optimizes the column configuration. If we were considering noninteger variables in the GA portion, we would need significantly more chromosomes. This can be seen by examining just the column length (L_C : 10–100). For this variable, if we consider an accuracy of $1.00\text{E-}2$ satisfactory, there are approximately $1.00\text{E+}4$ possibilities (this is already more than the total number of column configurations). If we were optimizing additional continuous variables by GA, it would become necessary to take the product for all of those to describe the complete search space. This example highlights the computational advantage of using SA to optimize the model's continuous variables.

SAGA employs both mutation and crossover procedures to perturb the current population to create a new, potentially better, population (better values for the objective function). These procedures essentially reward “genes” that are deemed to be beneficial, (i.e., the probability of removing a value for the number of columns in a zone decreases with uniformity of chromosomes). The new population is then checked against previous populations using a stagnation condition, which terminates the procedure when the fitness of the chromosomes is not seen to improve over a fixed period of time/generations. If the population meets the criteria, the current best chromosome from the population will be considered the optimizer, and an exhaustive SA procedure is performed to determine the optimal values of the other variables for this column configuration. Otherwise the evaluation process continues given this new population.

The ability of the SA algorithm to quickly and accurately determine a fitness value for a possible column configuration (chromosomes), given the sophistication level of this constrained optimization problem, is paramount to the success of this procedure. Implementation of SAGA was coded in MATLAB 7.0 and run on a 2.2 GHz Pentium 4. Table 4 summarizes the distribution of computation times using SAGA, for the solution of an optimization problem presented in the next section. As will be seen, the computational efficiency of SAGA allows us to undertake the analysis of a multi-objective optimization problem comprised of more than one hundred points in a few hours.

4 Results and discussion

In this section, the results of optimization by SAGA are presented and discussed for both single- and multi-objective optimizations. The results from two single-objective optimizations, maximum productivity and minimum cost will be compared with the results obtained for the same optimizations by use of the SWAT algorithm. It will be shown

that in each optimization the SAGA algorithm determines an SMB system that exhibits improved performance when compared to those determined by SWAT. The largest difference between the two algorithms, more than 9 percent, was observed for the cost minimization of the high-pressure system (5.2 MPa). These comparisons lead to the conclusion that the SAGA procedure offers additional robustness in determining the optimal SMB system. The results of a multi-objective optimization of maximum productivity and minimum desorbent consumption are presented and discussed in the last section.

The specific example to illustrate and evaluate the use of SAGA is the problem of the resolution of racemic mixtures of FTC-esters. This separation was originally examined by Xie et al. (2003a), where the reported stationary phase, mobile phase, isotherm parameters, and mass-transfer parameters appear in Table 1. More recently, an in-depth analysis of this separation problem was presented in Cauley et al. using SWAT optimization tool (Cauley et al. 2006), which will be used as a benchmark in this paper.

In each section three SMB systems are analyzed; low-pressure system (1.0 MPa), medium-pressure system (2.4 MPa), and high-pressure system (5.2 MPa). For each system the maximum operating pressure appears as the limit in the pressure drop constraint. In all cases, the switching time is set at 0.5 minute, yield is fixed at 0.99 for both enantiomers, and annual production of 25,000 kg is assumed. The bounds for the decision variables used in all optimizations are shown in Table 2, and the cost information used in Sect. 4.2 are shown in Table 3.

Table 1 Input parameters for the FTC separation system

Minimum step time $t_{s,\min}$ (min)	0.5	
Yield, Y (%)	99	
Purity, Pr (%)	99	
Interparticle void fraction, ε_b	0.24	
Intraparticle void fraction, ε_p	0.60	
Extra-column dead volume (% of the bed volume)	5.73	
Viscosity μ (g/cm/min)	0.32	
Mobile phase density ρ (g/cm ³)	0.79	
Packing density ρ_B (g/cm ³)	0.60	
Brownian diffusivity D_∞ (cm ² /min)	6.12×10^{-4}	
Intraparticle diffusivity D_p (cm ² /min)	2.0×10^{-5}	
k_f (cm/min)	Wilson and Geankoplis (1966) correlation	
E_b (cm ² /min)	Chung and Wen (1968) correlation	
Isotherms	(+)-FTC	$a = 1.12, b = 0.0168$
	(-)-FTC	$a = 2.12, b = 0.0383$

Table 2 Bounds of variables for optimal SMB design for FTC-ester separation system

Variable	Bounds
Feed concentration (g racemate/L) for FTC	10–80
CSP particle diameter d_p (μm)	5–200
Column diameter d_c (cm)	1–300
Column length L_c (cm)	10–100
Total number of columns N_{col}	8–20
Number of columns in each zone N^j	$2-(N_{col} - 6)$
Zone velocity u_0^j (cm/min)	0.01–200
Average port moving velocity v (cm/min)	0.01–20

Table 3 Costs and other information used in all optimizations

Production rate (kg/year)	25,000
CSP (Chiralpak AD) price (\$/kg) (CSP price is independent of the value of material parameters)	15,000
CSP depreciation period (years)	4
Solvent (methanol) price (\$/L)	0.20
Solvent recycle ratio (%)	95
Solvent recovery cost (\$/L)	0.10
Equipment depreciation period (years)	7
Acquisition price of SMB equipment	Low pressure system (≤ 1.0 MPa): \$700,000 Medium pressure system (≤ 2.4 MPa): \$1,000,000 High pressure system (≤ 13.8 MPa): is a function of column diameter d_c (cm), ($-927.67d_c^2 + 260,600d_c +$ $96,292$)

4.1 Single-objective optimizations: maximum productivity (PR)

In each case to maximize system productivity, SAGA selected a nine-column configuration, where Zone 3 contains three columns and the other zones contain two columns each. This is in contrast to the SWAT (Cauley et al. 2006) where identical or similar optimizations resulted in a configuration with two columns for each zone for a total of eight columns. While the differences in productivity between the eight- and nine-column systems are small, the more sophisticated search associated with the SAGA was able to identify an improvement in system design. In this case, the differences in system PR, about 3 percent, may or may not be economically important, but this result illustrates the robustness of SAGA to select an SMB system with higher PR and different column configuration that was overlooked by SWAT. Furthermore, given that the SA portion of the two algorithms use the same tolerance (i.e., $1.00\text{E-}8$ were used for all constraints), we can conclude that the differences observed are

attributable to differences in algorithms rather than differences in tolerances.

Table 4 shows the above results and the results of optimizations performed by SWAT (Cauley et al. 2006). In all optimizations, the optimal value for the variable feed concentration correspond to its upper bound value of 80 (g racemate/L), and the optimal value for the variable column length is its lower bound value of 10 cm. The optimal switching times to achieve maximum productivity are 2, 1.3, and 0.9 minutes, respectively, for 1.0, 2.4, and 5.2 MPa systems, were found to be above the minimum port-switching time of 0.5 min. The SAGA algorithm selected larger particle size, and higher interstitial velocity in all zones compared to the SWAT algorithm.

4.2 Single-objective optimizations: minimum average cost (AC)

For all cases the optimizations by the SAGA algorithm resulted in a lower AC compared to the SWAT algorithm. The highest difference between the two algorithms, more than 9 percent, was observed for the high-pressure system (5.2 MPa). Additionally, the SMB high-pressure system determined by SAGA has a very different column configuration, 2–5–7–2, as compared to the 2–2–2–2 column configuration found by the SWAT algorithm.

Table 5 summarizes the results of the two optimizations. As was observed in the previous section, the optimal value for the variable feed concentration correspond to its upper-bound value of 80 (g racemate/L) and, the optimal value for column length is its lower-bound value of 10 cm in all optimizations. In general, SAGA selected larger particles, longer columns, and increased the interstitial velocity in each zone as compared to SWAT.

4.3 Multi-objective optimization—maximum PR and minimum DC

For many real world problems, it is difficult to decide on a single objective. Often there are several competing objectives that are relevant and expected to be included in the decision process (Zhang et al. 2002, 2003; Cauley et al. 2004). Frequently, a Pareto curve (nondominated solutions) is developed to examine the trade-offs between objectives. Such a curve consists of the set of efficient points (Deb 2001). Specifically, these points represent the solutions (values) to the two objectives where any attempt to make one objective better will make the other objective worse. The two objectives considered in this section are: maximum productivity (PR) and minimum desorbent consumption (DC).

To construct the Pareto-efficient curve, the SWD model was augmented with an additional constraint requiring a lower bound (lb) on the value of PR: $PR \geq lb$. SAGA was

Table 4 Objective maximum productivity chiral separation of FTC for SMB with maximum operating pressures 1.0, 2.4, and 5.2 MPa

FTC	Objective maximum productivity (MPa) 1.0		Objective maximum productivity (MPa) 2.4		Objective maximum productivity (MPa) 5.2	
SMB equipment	SWAT	SAGA	SWAT	SAGA	SWAT	SAGA
Particle diameter (μm)	47.47	52.91	38.46	42.86	31.82	35.46
Column length (cm)	10.00	10.00	10.00	10.00	10.00	10.00
Column diameter (cm)	89.34	83.06	72.12	67.08	59.51	55.37
Total number of columns	8	9	8	9	8	9
Column configuration	2–2–2–2	2–2–3–2	2–2–2–2	2–2–3–2	2–2–2–2	2–2–3–2
Feed concentration (g racemate/L)	80	80	80	80	80	80
u_0^I (cm/min)	28.02	31.28	42.91	47.89	62.93	70.22
u_0^{II} (cm/min)	19.89	22.11	30.46	33.86	44.66	49.64
u_0^{III} (cm/min)	20.69	23.04	31.68	35.27	46.46	51.72
u_0^{IV} (cm/min)	18.98	20.80	29.07	31.86	42.63	46.71
Switching time (min)	2.14	1.95	1.40	1.27	0.96	0.87
Productivity (kg product/kg CSP/day)	0.2276	0.2341	0.3492	0.3589	0.5129	0.5268

Table 5 Objective minimum average cost-chiral separation of FTC for SMB with maximum operating pressures 1.0, 2.4, and 5.2 MPa

FTC	Objective minimum average cost (MPa) 1.0		Objective minimum average cost (MPa) 2.4		Objective minimum average cost (MPa) 5.2	
SMB equipment	SWAT	SAGA	SWAT	SAGA	SWAT	SAGA
Particle diameter (μm)	43.77	48.9	34.46	38.44	29.3	48.33
Column length (cm)	10.00	10.00	10.00	10.00	10.00	10.00
Column diameter (cm)	90.95	84.60	74.38	69.30	60.60	45.22
Total number of columns	8	9	8	9	8	16
Column configuration	2–2–2–2	2–2–3–2	2–2–2–2	2–2–3–2	2–2–2–2	2–5–7–2
Feed concentration (g racemate/L)	80	80	80	80	80	80
u_0^I (cm/min)	23.75	26.58	34.30	38.25	53.33	78.90
u_0^{II} (cm/min)	16.87	18.84	24.37	27.12	37.88	52.57
u_0^{III} (cm/min)	17.64	19.73	25.52	28.45	39.61	55.68
u_0^{IV} (cm/min)	16.22	17.93	23.50	25.92	36.45	47.95
Switching time (min)	2.51	2.27	1.73	1.57	1.12	0.81
Average Cost (\$/kg)	79.31	77.94	64.51	63.57	120.53	109.16

run multiple times, and for each, a new value of lb , from the feasible range for PR, was used in the above constraint. Figure 2 shows the values of the two objectives for the three pressure cases, where PR is on the x -axis and the minimum DC achieved for that PR is on the y -axis. Table 6 shows a summary of computation times to obtain each point plotted on the figure. The computational efficiency of SAGA is illustrated by a mean computation time of approximately 4.3 minutes for a point.

In all optimizations the optimal feed concentration is 80 g/L. In Fig. 2 we observe that for all cases the mini-

um desorbent consumption for achieving a desired level of PR increases, and the increase is exponential as the PR approaches its maximum level for the system.

Figures 3, 4, and 5 show some of the optimal decision variables. A close examination of graph A in the three figures show that SAGA finds increasing the particle size results in increasing PR, while keeping the DC at a minimum. However, the optimal particle size decreases as the maximum operating pressure increases. The optimal particle size to achieve maximum productivity is around 53, 43, and 35 μm respectively for 1.0, 2.4, and 5.2 MPa systems.

Fig. 2 Comparison of Pareto-efficient solutions for 1.0, 2.4, and 5.2 MPa maximum pressures; objectives maximum productivity and minimum desorbent consumption

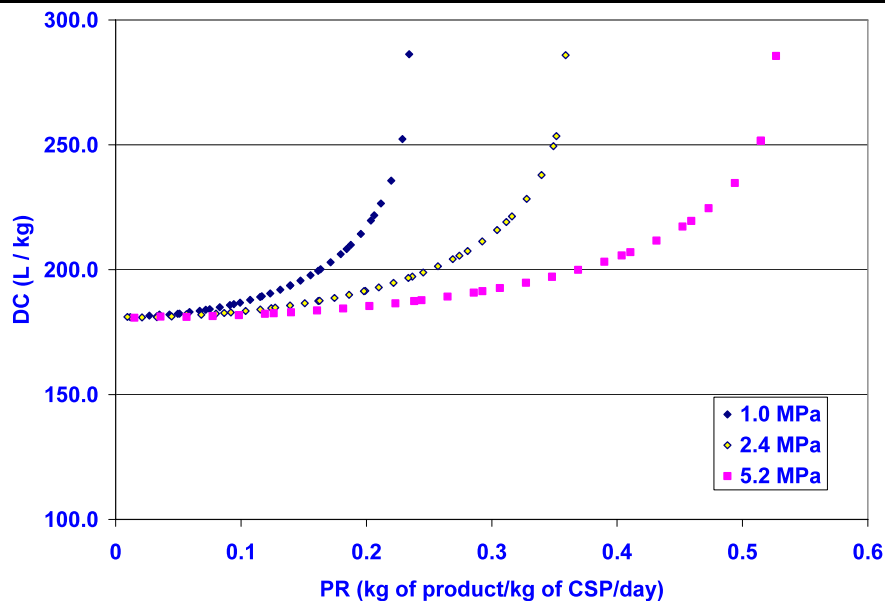


Table 6 Summary of SAGA computation time per optimization

Mean	4.29 min
Median	3.72 min
Standard Deviation	2.06 min
Minimum	1.90 min
Maximum	15.24 min
Sample size	116

Higher productivity is closely related to higher zone velocities and shorter switching times, as shown in graphs B, C, and D in the three figures. To increase PR while keeping the DC low, SAGA decreases the switching time as shown in graph B in the three systems. The optimal switching times to achieve maximum productivity are 2, 1.3, and 0.9 minutes, respectively, for 1.0, 2.4, and 5.2 MPa systems. The feed velocity, which is the difference between the velocities in Zone III and Zone II, is relatively low for all three systems in the range close to the point where maximum productivity is achieved.

SAGA selects the smallest diameter columns that can meet the production requirement in order to achieve the maximum productivity. As shown in graphs E and F in the three figures, SAGA chose as the optimal system, nine columns with the column length of 10 cm with the exception of points that correspond to very low productivity. The column length of 10 cm is the lower bound used in this study. A shorter column permits the use of smaller particle size, which results in higher PR. The choice of the nine and more columns for an SMB system is unlike what is reported in the literature for this problem (Lee et al. 2005b; Cauley et al. 2006).

5 Conclusion

Earlier research has concluded that the stochastic optimization algorithm, SWAT, was at least as accurate and in order of magnitude more computationally efficient than a deterministic sampling procedure (grid search) in determining the optimal design of a SMB system. A potential source of error in SWAT was the use of a deterministic branch-and-bound algorithm to determine the system configuration.

This paper introduces a new algorithm that is a hybrid between two stochastic optimizing techniques, GA and SA. By replacing the deterministic branch and bound in SWAT with the stochastic GA for determining the integer variables of the model, we are presenting an algorithm that is both robust and computationally efficient. SAGA utilizes the best of SA and GA for optimizing the design of SMB systems. Results of single- and multi-objective optimizations are presented and discussed.

Nomenclature

- a_i = Langmuir isotherm parameter of component i based on solid volume, L/L S.V.
- AC = average purification cost, \$/kg
- b_i = Langmuir isotherm parameter of component i , L/g
- CSPC = average CSP cost, \$/kg
- C_{CSP} = CSP price, \$/kg
- C_E = acquisition price of SMB equipment, \$
- $C_{F,i}$ = feed concentration of component i , g/L
- $C_{p,i}$ = plateau concentration of component i , g/L
- $C_{s,i}$ = diluted plateau concentration of component i at the feed port, g/L

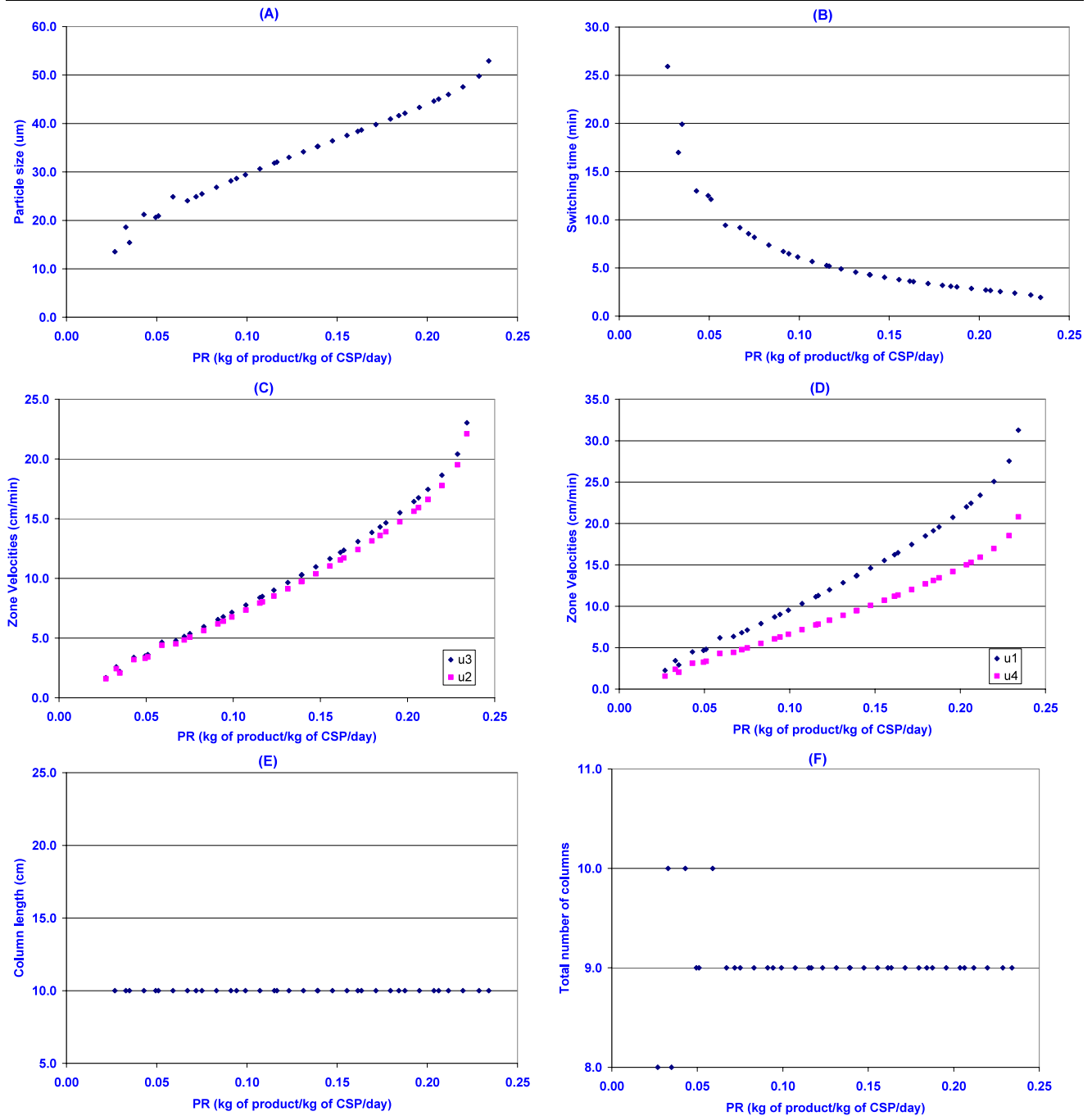


Fig. 3 Pareto-efficient solutions for decision variables for 1.0 MPa maximum pressure; objectives maximum productivity and minimum desorbent consumption. The optimal feed concentration is 80 g/L for all solutions and yield is fixed at 0.99 for both enantiomers

C_{Sol} = solvent price, \$/L

C_{sol}^{Rcy} = solvent recovery price, \$/L

d_c = column diameter, cm

d_p = particle diameter, μm

$D_{P,i}$ = intra-particle diffusivity of component i , cm^2/min

$D_{\infty,i}$ = Brownian diffusivity of component i , cm^2/min

DC = desorbent consumption, L/kg

DV = extra-column dead volume, cm^3

EqC = average equipment cost, \$/kg

$E_{b,i}^j$ = axial dispersion coefficient of component i in zone j , cm^2/min

$k_{f,i}^j$ = film mass transfer coefficient of component i in zone j , cm/min

$K_{e,i}$ = size exclusion factor for component i

K_i^j = lumped mass transfer coefficient of component i in zone j , min^{-1}

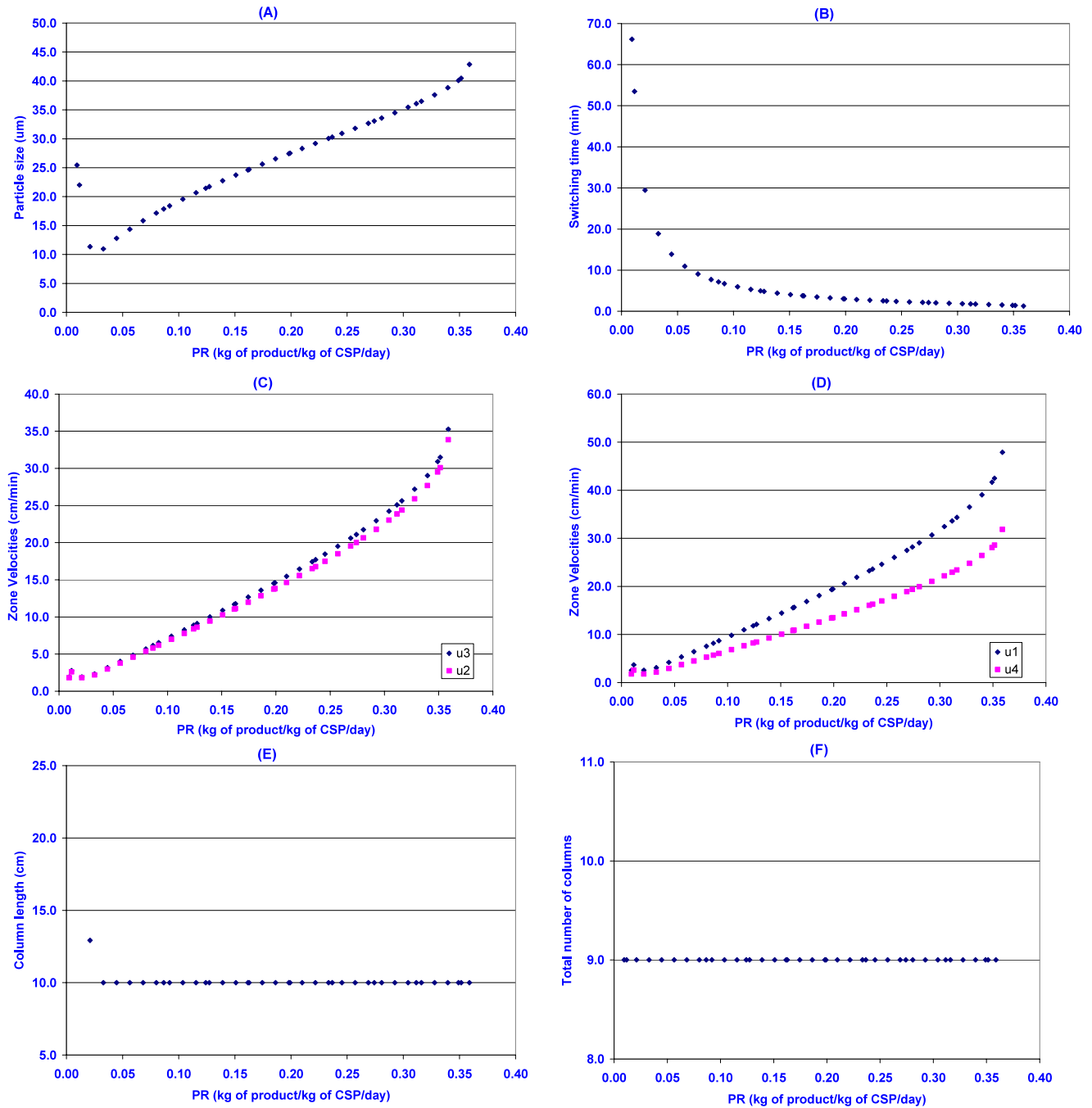


Fig. 4 Pareto-efficient solutions for decision variables for 2.4 MPa maximum pressure; objectives maximum productivity and minimum desorbent consumption. The optimal feed concentration is 80 g/L for all solutions and yield is fixed at 0.99 for both enantiomers

L_c = single column length, cm
 N_{col} = total number of columns in SMB
 N_{col}^j = number of columns in zone j
 $P = \frac{1-\varepsilon_b}{\varepsilon_b}$ = Phase ratio
 PR = productivity, kg product/ kg CSP/ day
 Pur_{req}^E = purity required for extract
 Pur_{req}^R = purity required for raffinate
 Q = annual production, g

r_E = depreciation period of equipment, year
 r_{CSP} = depreciation period of CSP, year
 S = column cross sectional area, cm^2
 $SolC$ = average solvent cost, \$/kg
 S_{Rcy} = solvent recycle ratio, %
 t = duration of production, min
 $t_{s,min}$ = minimum step time, min
 $u_{F,max}$ = maximum feed interstitial velocity, cm/min

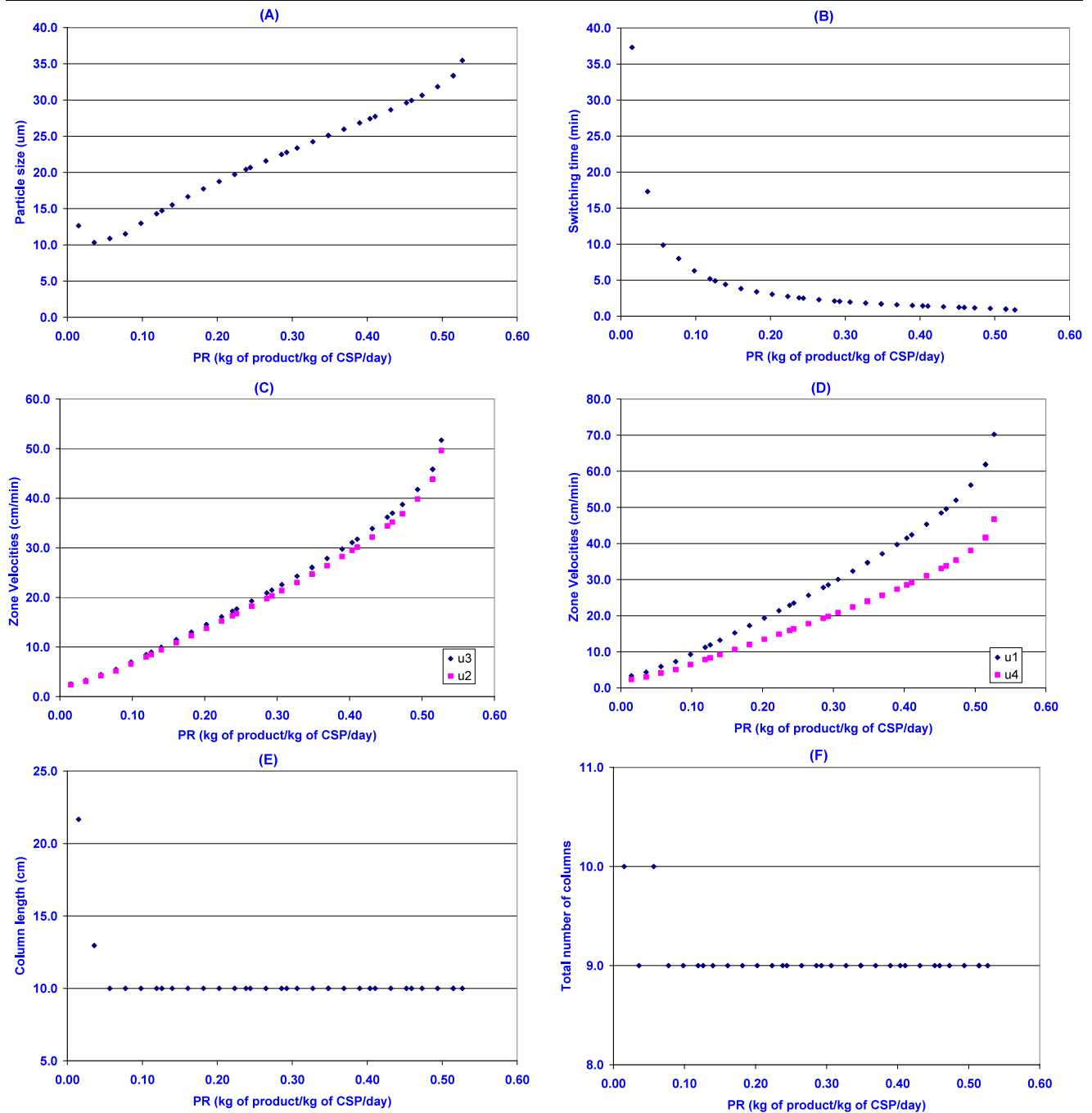


Fig. 5 Pareto-efficient solutions for decision variables for 5.2 MPa maximum pressure; objectives maximum productivity and minimum desorbent consumption. The optimal feed concentration is 80 g/L for all solutions and yield is fixed at 0.99 for both enantiomers

u_0^j = interstitial velocity in zone j , cm/min
 v = average port moving velocity, cm/min
 Y_i = yield of component i

Greek Letters

β_i^j = decay factor of standing component i in zone j
 δ_i = retention factor of component i

Δp^j = pressure drop for zone j , MPa
 Δp_{\max} = maximum pressure drop, MPa
 ε_b = interparticle void fraction
 ε_p = intraparticle void fraction (or particle porosity)
 ρ = mobile phase density, g/mL
 ρ_B = bulk packing density, kg/m³
 μ = viscosity, g/cm/s

Subscripts and Superscripts

i = component index
 j = zone number index

Acknowledgement The first author was a visiting scholar at the School of Chemical Engineering, Purdue University. Support from Seaver Research Council Grant at Pepperdine University and Indiana 21st Century Research are gratefully acknowledged. The authors are thankful to Mr. Geoffrey B. Cox for insightful discussions and comments, and valuable comments from two anonymous reviewers.

Appendix: Model of nonideal nonlinear SMB system for binary separation

Input parameters: Pur_{req}^E , Pur_{req}^R , $DV'(\frac{DV}{L_c S})$, $t_{s,min}$, Δp_{max} , $\varepsilon_b(P = \frac{1-\varepsilon_b}{\varepsilon_p})$, ε_p , ρ , ρ_B , μ , and a_i , b_i , $D_{p,i}$, $D_{\infty,i}$, $K_{e,i}$, Y_i for $i = 1, 2$

Decision variables:

Tier0 variables: $C_{F,1}$, $C_{F,2}$.

Tier1 variables: S , d_p , L_c , N_{col} .

Tier2 variables: N_{col}^j , $j = I, II, III$.

Tier3 variables: v , u_0^j , $j = I, II, III, IV$, $C_{s,i}$, $i = 1, 2$

Derived variables:

$$\lambda_F = 4a_1a_2b_1b_2C_{F,1}C_{F,2};$$

$$\chi_F = -a_1(1 + b_2C_{F,2}) + a_2(1 + b_1C_{F,1});$$

$$\gamma_{+,F} = \frac{\chi_F + \sqrt{\chi_F^2 + \lambda_F}}{2a_2b_1C_{F,2}}; \quad \lambda = 4a_1a_2b_1b_2C_{s,1}C_{s,2};$$

$$\chi = -a_1(1 + b_2C_{s,2}) + a_2(1 + b_1C_{s,1});$$

$$\gamma_- = \frac{\chi - \sqrt{\chi^2 + \lambda}}{2a_2b_1C_{s,2}};$$

$$C_{p,1} = \frac{a_1 - a_2}{a_2b_1 + \frac{a_1b_2}{\gamma_-}}; \quad C_{p,2} = \frac{a_2 - a_1}{a_1b_2 + a_2b_1\gamma_{+,F}};$$

$$\delta_2^I = \varepsilon_p + (1 - \varepsilon_p)a_2 + \frac{DV'}{1 - \varepsilon_b};$$

$$\delta_1^{II} = \varepsilon_p + (1 - \varepsilon_p)\left(\frac{a_1}{1 + b_2C_{p,2}}\right) + \frac{DV'}{1 - \varepsilon_b};$$

$$\delta_2^{III} = \varepsilon_p + (1 - \varepsilon_p)\left(\frac{a_2}{1 + b_1C_{s,1} + b_2C_{s,2}}\right) + \frac{DV'}{1 - \varepsilon_b};$$

$$\delta_1^{IV} = \varepsilon_p + (1 - \varepsilon_p)\left(\frac{a_1}{1 + b_1C_{p,1}}\right) + \frac{DV'}{1 - \varepsilon_b};$$

Axial dispersion coefficients:

$$E_b^j = \frac{u_0^j \varepsilon_b (d_p 10^{-4})}{0.2 + 0.011 \left(\frac{\rho u_0^j \varepsilon_b d_p 10^{-4}}{60\mu} \right)^{0.48}}, \quad \text{for } j = I, II, III, IV$$

Film mass transfer coefficients:

$$k_{f,i}^j = 1.09(u_0^j)^{\frac{1}{3}} \left(\frac{\varepsilon_b d_p 10^{-4}}{D_{\infty,i}} \right)^{-\frac{2}{3}},$$

where for $j = I, III$, $i = 2$ and for $j = II, IV$, $i = 1$

Overall mass transfer coefficients:

$$\frac{1}{K_2^j} = \frac{(d_p 10^{-4})^2}{60K_{e,i} \varepsilon_p D_{p,i}} + \frac{d_p 10^{-4}}{6k_{f,i}^j},$$

where for $j = I, III$, $i = 2$ and for $j = II, IV$, $i = 1$

Ratio of highest concentration to the lowest concentration of the standing wave in each zone:

$$\beta_2^I = \ln \left(\frac{u_0^I(u_0^{III} - u_0^{IV})Y_2}{u_0^{IV}(u_0^I - u_0^{II})(1 - Y_2)} \right);$$

$$\beta_1^{II} = \ln \left(\frac{(u_0^I - u_0^{II})(u_0^{III}C_{s,1} - (u_0^{III} - u_0^{II})C_{F,1})}{u_0^{II}(u_0^{III} - u_0^{II})C_{F,1}(1 - Y_1)} \right)$$

$$\beta_2^{III} = \ln \left(\frac{(u_0^{III} - u_0^{IV})C_{s,2}}{(u_0^{III} - u_0^{II})C_{F,2}(1 - Y_2)} \right);$$

$$\beta_1^{IV} = \ln \left(\frac{u_0^{IV}(u_0^I - u_0^{II})Y_1}{u_0^I(u_0^{III} - u_0^{IV})(1 - Y_1)} \right)$$

Maximum feed interstitial velocity:

$$u_{F,max} = \frac{P^2(\delta_2^{III} - \delta_1^{II})^2}{4 \left(\frac{P\beta_2^{III}(\delta_1^{III})^2}{K_2^{III}N_{col}^{III}L_c} + \frac{P\beta_1^{II}(\delta_1^{II})^2}{K_1^{II}N_{col}^{II}L_c} \right)} - \left(\frac{\beta_2^{III}E_b^{III}}{N_{col}^{III}L_c} + \frac{\beta_1^{II}E_b^{II}}{N_{col}^{II}L_c} \right)$$

Subject to constraints:

I. Bounds on variables:

$$C_{F,min} \leq C_{F,1} + C_{F,2} \leq C_{F,max}; \quad S_{min} \leq S \leq S_{max};$$

$$d_{p,min} \leq d_p \leq d_{p,max}; \quad L_{c,min} \leq L_c \leq L_{c,max};$$

$$N_{col,min} \leq N_{col} \leq N_{col,max}; \quad N_{col}^j \geq 2, \quad j = I, II, III;$$

$$u_{min}^j \leq u_0^j \leq u_{max}^j,$$

$$j = I, II, III, IV; \quad v \leq \frac{L_{c,min}}{t_{s,min}}$$

II. Linear system constraints:

$$C_{s,i} \leq C_{F,i}, \quad i = 1, 2; \quad C_{F,1} = C_{F,2};$$

$$u_0^{III} < u_0^I; \quad u_0^{II} < u_0^{III}; \quad u_0^{IV} < u_0^I;$$

$$u_0^{II} < u_0^I; \quad u_0^{IV} < u_0^{III}$$

III. Non-linear system constraints:

$$(a_2 b_1 \gamma_{+,F}^2 + a_1 b_2 \gamma_{+,F}) C_{s,2} - (a_2 b_1 \gamma_{+,F} + a_1 b_2) C_{s,1} + (a_1 - a_2) \gamma_{+,F} = 0$$

$$u_0^{\text{III}} - u_0^{\text{II}} \leq u_{F,\max}; \quad u_0^{\text{III}} C_{s,1} - (u_0^{\text{III}} - u_0^{\text{II}}) C_{F,1} > 0;$$

$$\frac{u_0^{\text{III}} - u_0^{\text{II}}}{u_0^{\text{III}} - u_0^{\text{IV}}} C_{F,1} Y_1 = C_{p,1};$$

$$Pur^E = \frac{C_{F,2} Y_2}{C_{F,2} Y_2 + C_{F,1} (1 - Y_1)} \geq Pur_{req}^E;$$

$$Pur^R = \frac{C_{F,1} Y_1}{C_{F,1} Y_1 + C_{F,2} (1 - Y_2)} \geq Pur_{req}^R$$

IV. SW equations:

$$v - \frac{u_0^{\text{I}}}{1 + P \delta_2^{\text{I}}} = - \frac{\beta_2^{\text{I}}}{(1 + p \delta_2^{\text{I}}) N_{\text{col}}^{\text{I}} L_c} \left(E_b^{\text{I}} + \frac{P v^2 (\delta_2^{\text{I}})^2}{K_2^{\text{I}}} \right)$$

$$v - \frac{u_0^{\text{II}}}{1 + P \delta_1^{\text{II}}} = - \frac{\beta_1^{\text{II}}}{(1 + p \delta_1^{\text{II}}) N_{\text{col}}^{\text{II}} L_c} \left(E_b^{\text{II}} + \frac{P v^2 (\delta_1^{\text{II}})^2}{K_1^{\text{II}}} \right)$$

$$v - \frac{u_0^{\text{III}}}{1 + P \delta_3^{\text{III}}} = - \frac{\beta_3^{\text{III}}}{(1 + p \delta_3^{\text{III}}) N_{\text{col}}^{\text{III}} L_c} \left(E_b^{\text{III}} + \frac{P v^2 (\delta_3^{\text{III}})^2}{K_3^{\text{III}}} \right)$$

$$v - \frac{u_0^{\text{IV}}}{1 + P \delta_4^{\text{IV}}} = - \frac{\beta_4^{\text{IV}}}{(1 + p \delta_4^{\text{IV}}) N_{\text{col}}^{\text{IV}} L_c} \left(E_b^{\text{IV}} + \frac{P v^2 (\delta_4^{\text{IV}})^2}{K_4^{\text{IV}}} \right)$$

$$L_c - v t_{s,\min} \geq 0$$

V. Demand constraint:

$$Q_{\min} \leq Q \leq Q_{\max}; \quad \text{for } Q = \frac{1}{10^3} S \varepsilon_b (u_0^{\text{III}} - u_0^{\text{II}}) C_{F,p} Y_p t$$

VI. Pressure drop constraint: $\sum_j \Delta p^j = \Delta p_{\max}$; where

$$\Delta p^j = \left(\frac{150 \mu u_0^j N_{\text{col}}^j L_c P^2}{d_p^2} \left(\frac{10^6}{6} \right) + \frac{1.75 \rho (u_0^j)^2 N_{\text{col}}^j L_c P}{d_p} \left(\frac{1}{3.6} \right) \right) 10^{-6},$$

$$j = \text{I, II, III, IV}$$

Objective functions used in the optimizations:

$$\text{Productivity (PR): } PR = \frac{Q}{\rho_B V (365)}$$

$$\text{Desorbent Consumption (DC): } DC = \frac{1000(u^{\text{I}} - u^{\text{IV}})}{(u^{\text{III}} - u^{\text{II}}) C_{F,p} Y_p}$$

$$\text{Average Cost (AC): } AC = EqC + CSPC + SolC \text{ (\$/kg),}$$

where

$$EqC = \frac{C_E}{Qr_E}; \quad CSPC = \frac{C_{CSP} S L_c \rho_B N_{\text{col}}}{Qr_{CSP}};$$

$$SolC = \frac{S \varepsilon_b t}{Q} [(u_0^{\text{I}} - u_0^{\text{IV}}) + (u_0^{\text{III}} - u_0^{\text{II}})] \\ \times [C_{sol} (1 - S_{Rcy}) + C_{sol}^{Rcy} S_{Rcy}]$$

References

- Biegler, L.T., Grossmann, I.E.: Retrospective on optimization. *Comput. Chem. Eng.* **28**, 1169–1192 (2004)
- Broughton, D.B.: Production-scale adsorptive separations of liquid mixtures by simulated moving-bed technology. *Sep. Sci. Prog.* **66**, 70 (1961S)
- Cauley, F.G., Xie, Y., Wang, N.-H.L.: Optimization of SMB systems with linear adsorption isotherms by the standing wave annealing technique. *Ind. Eng. Chem. Res.* **43**, 7588–7599 (2004)
- Cauley, F.G., Cauley, S.F., Lee, K.B., Xie, Y., Wang, N.-H.L.: Standing wave annealing technique: for the design and optimization of nonlinear simulated moving bed systems with significant mass-transfer effects. *Ind. Eng. Chem. Res.* **45**, 8697–8712 (2006)
- Cerny, V.: Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. *J. Optim. Theory Appl.* **45**, 41–51 (1985)
- Chung, S.F., Wen, C.Y.: Longitudinal dispersion of liquid flowing through fixed and fluidized beds. *AIChE J.* **14**, 857 (1968)
- Davis, L.: Genetic Algorithms and Simulated Annealing. Pitman, London (1987)
- Davis, L.: Hand Book of Genetic Algorithms. Van Nostrand Reinhold, New York (1991)
- Deb, K.: Multi-Objective Optimization using Evolutionary Algorithms. Wiley, New York (2001)
- Holland, J.H.: Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications of Biology, Control and Artificial Intelligence. University of Michigan Press, Ann Arbor (1975)
- Kirkpatrick, S., Gelatt, C.D. Jr., Vecchi, M.P.: Optimization by simulated annealing. *Science* **220**, 671–680 (1983)
- Lee, K.B., Chin, C.Y., Xie, Y., Cox, G., Wang, N.-H.L.: Standing wave design of a simulated moving bed under a pressure limit for enantioseparation of phenylpropanolamine. *Ind. Eng. Chem. Res.* **44**, 3249–3267 (2005a)
- Lee, K.B., Mun, S., Cauley, F.G., Cox, G., Wang, N.-H.L.: Optimal standing-wave design of nonlinear simulated moving bed systems for enantioseparation. *Ind. Eng. Chem. Res.* **45**, 739–752 (2005b)
- Ludemann-Hombourger, O., Bailly, M., Nicoud, R.M.: Design of a simulated moving bed: optimal particle size of the stationary phase. *Sep. Sci. Technol.* **35**, 1285–1305 (2000)
- Luenberger, D.G.: Linear and Nonlinear Programming, 2nd edn. Kluwer Academic, Norwell (2003)
- Ma, Z., Wang, N.-H.L.: Standing wave analysis of SMB chromatography: linear systems. *AIChE J.* **43**, 2488–2508 (1997)
- Mallmann, T., Burris, B.D., Ma, Z., Wang, N.-H.L.: Standing wave design of nonlinear SMB systems for fructose purification. *AIChE J.* **44**, 2628–2646 (1998)
- Mazzotti, M., Storti, G., Morbidelli, M.: Optimal operation of simulated moving bed units for nonlinear chromatographic separations. *J. Chromatogr. A* **769**, 3–24 (1997)
- Michalewicz, Z., Fogel, D.B.: How to Solve It: Modern Heuristics. Springer, New York (2000)
- Mun, S., Xie, Y., Kim, J.-H., Wang, N.-H.L.: Optimal design of a size-exclusion tandem simulated moving bed process for insulin purification. *Ind. Eng. Chem. Res.* **42**, 1977–1993 (2003)
- Ruthven, D.M., Ching, C.B.: Counter-current and simulated counter-current adsorption separation processes. *Chem. Eng. Sci.* **44**, 1011–1038 (1989)
- Salamon, P., Sibani, P., Frost, R.: Facts, Conjectures, and Improvements for Simulated Annealing. SIAM, Philadelphia (2002)
- Schulte, M., Strube, J.: Preparative enantioseparation by simulated moving bed chromatography. *J. Chromatogr. A* **906**, 399–416 (2001)
- Silva, V.M.T., Minceva, M., Rodrigues, A.E.: Novel analytical solution for a simulated moving bed in the presence of mass-transfer resistance. *Ind. Eng. Chem. Res.* **43**, 4494–4502 (2004)

- Wales, D.J., Scherega, H.A.: Chemistry: global optimization of clusters, crystals and biomolecules. *Science* **27**, 1368–1372 (1999)
- Wilson, E.J., Geankoplis, C.J.: Liquid mass transfer at very low Reynolds numbers in packed beds. *Ind. Eng. Chem. Fundam.* **5**, 9 (1966)
- Wu, D.-J., Xie, Y., Ma, Z., Wang, N.-H.L.: Design of simulated moving bed chromatography for amino acid separations. *Ind. Eng. Chem. Res.* **37**, 4023–4035 (1998)
- Xie, Y., Koo, Y.-M., Wang, N.-H.L.: Preparative chromatographic separation: simulated moving bed and modified chromatography methods. *Biotechnol. Bioprocess. Eng.* **6**, 1 (2001)
- Xie, Y., Mun, S.-Y., Kim, J.H., Wang, N.-H.L.: Standing wave design and experimental validation of a tandem simulated moving bed for insulin purification. *Biotechnol. Prog.* **18**, 1332–1344 (2002)
- Xie, Y., Hritzko, B., Chin, C.Y., Wang, N.-H.L.: Separation of FTC-ester enantiomers using a simulated moving bed. *Ind. Eng. Chem. Res.* **42**, 4055–4067 (2003a)
- Xie, Y., Farrenburg, C., Chin, C.Y., Mun, S., Wang, N.-H.L.: Design of SMB for a nonlinear amino acid system with mass-transfer effects. *AIChE J.* **49**, 2850–2863 (2003b)
- Zhang, Z., Hidajat, K., Ray, A.K., Morbidelli, M.: Multiobjective optimization of SMB and varicol process for chiral separation. *AIChE J.* **48**, 2800–2816 (2002)
- Zhang, Z., Mazzotti, M., Morbidelli, M.: Multiobjective optimization of simulated moving bed and varicol processes using a genetic algorithm. *J. Chromatogr. A* **989**, 95–108 (2003)