Multiobjective Molecular Design for Integrated Process-Solvent Systems Synthesis

Athanasios I. Papadopoulos and Patrick Linke
Centre for Process and Information Systems Engineering, School of Engineering University of Surrey, Guildford, Surrey, GU2 7XH, U.K.

DOI 10.1002/aic.10715
Published online November 4, 2005 in Wiley InterScience (www.interscience.wiley.com).

The limitations are addressed that emerge from the computer aided molecular design (CAMD) applications in the optimal synthesis of solvents through a novel approach of designing solvents built upon multiple objective optimization technology. This technology has certain advantages over single objective optimization approaches that have been used for the design of optimal solvent molecules in the past. The different synthesis drives, described by the various objective functions, are treated independently without being submitted to unnecessary fixed assumptions that bias the designs. As a result, all the underlying trends and trade-offs among the properties of the candidate optimal molecules, as well as the structure – property relations are allowed to be revealed. The ability of the presented synthesis framework to embed highly inclusive solutions is elaborated, and the trends are revealed that dominate the relations between the utilized objective functions. It is also shown that solvent molecules, designed in the event that process performance feedback is accounted for, are existent in the Pareto optimal front. The proposed methodology is illustrated with two comprehensive examples in the design of solvents for liquid-liquid extraction and gas-absorption processes. © 2005 American Institute of Chemical Engineers *AIChE J,* 52: 1057–1070, 2005
Keywords: multiobjective optimization, computer aided solvent design, simulated annealing, liquid-liquid extraction, gas-absorption

Introduction

The increasing demand for solvent molecules in the process industry requires the design and synthesis of compounds to be used in high performance processes with a minimum environmental impact. Modern computer aided molecular design (CAMD) technology combines group contribution methods (GC) with knowledge-based or optimization-based technology in order to screen and scope for molecules with the desired properties. Molecules are generated through structural combinations of functional groups, and properties are predicted using thermodynamic models, such as UNIFAC. GC methods offer predictive capabilities at low computational cost. Knowledge-based approaches,[1,2,3,4,5,6] capitalize on knowledge introduced into expert models and are likely to produce designs that bear similarities with conventional, existing molecules without necessarily representing optimal choices. In contrast, the combined use of GC with optimization-based methods can be employed to identify both novel and existing molecules with the "best" performance. The generated molecules are designed for weighted optimality in a number of thermodynamic properties that are anticipated to have a significant effect on the performance of the process in which they will be utilized.

To date, a number of optimization-based CAMD methods have been proposed and have found a number of diverse applications. Deterministic local optimization-based CAMD methods have found many applications in areas, such as polymer design,[7] refrigerant design[8,9,10] and solvent design.[11,12,13,14,15,16] As an alternative philosophy, stochastic optimization technology has been adopted as it promises

Correspondence concerning this article should be addressed to P. Linke at p.linke@surrey.ac.uk.

potentially robust alternatives to identify solutions in the globally optimal domain. Stochastic optimization approaches include applications of Genetic Algorithms[17,18] and Simulated Annealing.[19,20,21,22] Poor search speeds and low success rates have been reported for GA applications in polymer design.[17] Van Dyk and Nieuwoudt[18] report successful results for the design of solvents for extractive distillation, but they are obtained for a limited search space. Marcoulaki and Kokossis[19,20] introduce a conceptual design framework that makes use of Simulated Annealing in their attempt to tackle the combinatorial complexity of the CAMD problem. They reported their approach to be robust and quick. As a result, there is no need to limit the number of groups or to introduce other constraints to limit the search space should this not be required due to a lack of reliability of the group contribution method used. This makes the approach unbiased toward the selection of "reasonable" groups. The applications in solvent selection for liquid-liquid extraction and extractive distillation reported by Marcoulaki and Kokossis[20] compare favorably to previous developments.

Advances in CAMD methods for solvents have mainly focussed on the development of optimization schemes and solvent design representations. However, there are a number of aspects that need to be addressed to enable highly successful CAMD applications in solvent design for optimal solvent-based process systems. The formulation of the objective function and the thermodynamic property constraints at the CAMD optimization stage has a crucial impact on the optimal solvent molecule identified, and, thus, on the performance that can be attained by the solvent-based process. The CAMD formulations need to accommodate for the optimization of a number of properties that will have an impact on the performance of the process in which the designed solvent molecule will be utilized. For instance, a desirable solvent should have a high selectivity, a low-heat of vaporisation, a high-relative volatility, and a boiling point that would allow solvent regeneration at reasonable temperatures. Since only one objective function is formulated in the existing CAMD methods, assumptions have to be made regarding the weighting of the individual property objectives. A different weighting of objectives will lead to the identification of different "optimal" solvent molecules, each of which would result in a different process performance in terms of process cost or environmental impact. As there is no clear link between property objectives and process performance, the assumption behind the CAMD objective functions and constraints are likely to lead to suboptimal performances of the overall process-solvent system. This is because important thermodynamic property effects can easily be excluded, over—, or underestimated as process performance feedback is unavailable at the CAMD stage. Without such feedback, appropriate levels of importance of the individual thermodynamic properties used to assess the molecular performance are extremely likely to be misrepresented using a single objective function combined with property constraints, simply because their impact on the process is unknown. In existing solvent design CAMD formulations, one selected thermodynamic property generally assumes primary importance after assigning the objective function attribute to it, restricting other relevant properties to an auxiliary role.

Apart from the uncertainty regarding the process impact, there is another limitation of current optimization-based CAMD methods. The final optimal solvent molecule(s) represent(s) just one (or a few) possible design instance(s) derived from a single design objective. This means that the underlying trade-offs and trends among the different relevant solvent properties, as well as the molecular structure-property relations are not revealed. Such information could be a valuable asset for the design of processes, based on the understanding of property and structural interactions. To this end, Marcoulaki and Kokossis[19,20] develop targets for the properties employed in the design and, thus, reveal part of this information. However, most of this information remains unknown.

Recognizing the importance of process performance feedback, a number of approaches have been reported recently that enable process performance feedback in molecular design.[6,13,14,15,16,20,22,23,24,25,26] Two distinct design philosophies were followed in these contributions. One design philosophy involves screening for solvent molecules based on prespecified property targets. The molecules meeting these targets are further screened with regards to their process performance either by participating in process simulation or process optimization. In these cases subsystems of the overall system are targeted, thus, the results are liable to assumptions regarding the size of the solvent-process design space. A second design philosophy is the simultaneous optimization of solvent-process supermodels which may involve separation and reactive separation processes. The enormous amount of solvent-process design options, and the complexities caused by the nonconvexities of the employed models pose challenges beyond the scope of current optimization technology and, therefore, introduce low confidence with regards to the quality of the obtained results.

In view of these problems, there is a need to develop a new strategy that allows the reliable and quick optimization of process-solvent systems. This work proposes an approach that introduces a multiobjective CAMD to decompose the solvent-process design problem. The resulting problem maintains the complexity at levels associated with the traditional CAMD approaches followed by process synthesis, while realizing the benefits of the most general integrated solvent-process design methods proposed to date. This article introduces the decomposition strategy and maintains a focus on CAMD using multiobjective optimization. A further publication will present in detail the overall design approach, implementation issues and applications in the design of a number of solvent-based process systems.

Problem decomposition for integrated process-solvent design

The above discussion has highlighted the problems resulting from the formulation of a single objective CAMD optimization problem as this introduces a high risk of failure to provide optimal solutions to the overall process-solvent design problem. Such failure is a direct result of the assumptions that have to be made to combine the different conflicting CAMD design objectives into a single objective function and a number of property constraints without knowledge about the effect these assumptions have on the performance of the process in which the solvent is to be used. By adopting such an approach, a large number of solvent candidates are excluded before process performance information is available, that is, in the process synthesis stage.

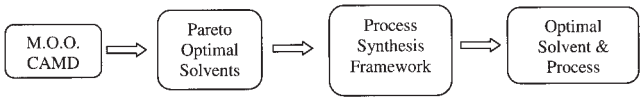

Figure 1. Proposed decomposition based strategy.

Instead, the formulation of the CAMD problem as a multiobjective optimization problem is proposed in order to capture the relations between all the physical properties and performance/environmental indices expected to impact on process performance and to extract solvent design candidates across the Pareto optimal front. This avoids the assignment of fixed, user-defined weights to the indices and properties and, thus, allows to carry over a set of potentially optimal solvent candidates to the process synthesis stage, one of which will be the optimal solvent for a given process performance indicator, such as minimum total annualized cost.

The proposed approach is illustrated in Figure 1. The first stage involves the solvent design level. The multiple objective optimization (MOO) methodology extracts the Pareto optimal solvent design candidates for the set of molecular design objectives important to the particular separation task under consideration. The optimization reveals property-property and property-structure relations among the solvent molecules independently of presumed process indices that would bias design configurations toward seemingly "optimal" directions. This information can be mapped onto the process synthesis problem which can then be optimized to yield the optimal solvent-process system. As a set of Pareto optimal solvents is developed using the proposed approach rather than a single optimal solvent design, the risk of potentially useful options being excluded from the CAMD solution space is greatly reduced. The Pareto front describes the optimal solvent design space for a given system of components and separation task regardless of process design and economics. Moreover, the problem complexity is kept at a moderate, that is, "solvable" level without discarding important design information, neither at the CAMD nor at the process synthesis end. Efficient ways of introducing the Pareto information into the process synthesis problem will be presented in a separate publication together with applications of the overall framework to solvent-based process design. This article focuses on the issues related to multiobjective objective optimization in CAMD.

The proposed approach is a generic solvent design strategy that allows the design of solvent molecules independently of process performance criteria and economics, whereas simultaneously maintaining inclusiveness of relevant solvent design options and facilitating the combinatorial complexity of the integrated solvent and process design problem. In this context it is important to demonstrate the ability of the presented synthesis methodology to embed highly inclusive solutions, to reveal the hidden trends and trade-offs between the different objectives and to exemplify its flexibility of application.

Optimization framework

This work addresses the synthesis of solvent molecules as a MOO. problem. The representation and optimization framework of the solvent molecules is based on the work by Marcoulaki and Kokossis.[19] All the attributes of the molecules regarding their representation during the simulation phase and perturbations during the optimization search follows the description given by Marcoulaki and Kokossis.[19] However, the acceptance criteria employed by simulated annealing are adjusted to the requirements of the MOO principles, and a stage where the candidate optimal solvent molecules are assessed is added.

Representation of solvent molecules

A brief description of the representation of the solvent molecules suggested by Marcoulaki and Kokossis[19] is presented in this section. Each molecule is represented by a molecular vector M_k composed by the UNIFAC group vector $m_{k,}$ and the composition matrix A_k. It becomes evident that unrestricted combinations of group vectors with composition matrices can lead to meaningless molecular vectors unless additional constraints are imposed to account for the groups, the valence, the class and the population of groups in the molecule. After the feasibility of the molecular vector M_k is confirmed perturbation moves take place. The moves consist of replacing groups that constitute the vector with alternative groups (substitution moves) and of augmenting or reducing the dimension K of the vector (expansion/contraction moves). As soon as the molecule receives its final form its required properties are calculated, and it is subjected to performance tests dictated by Simulated Annealing.

Multiple objective optimization

Multiple objective optimization is a technology that deals with the simultaneous optimization of more than one objective. Due to its ability to identify the diverse trends involved in each objective, thus, providing desired design insights, it is widely used in the chemical engineering domain. Applications of multiple objective optimization in chemical engineering problems involve environmentally friendly utility design,[27] waste minimization,[28] polymerization reactor design[29] and many others.

In mathematical terms a MOO problem is formulated as follows:

Let $f_i(x)(i=1\cdots p)$: $D \rightarrow R$ be a set of p objective functions to be optimized, where x is a feasible solution, and D is a set of feasible solutions

$$\text{Optimize } f_1(x), f_2(x), \ldots f_p(x) \tag{1}$$

$$\text{subject to } x \in D \tag{2}$$

Since usually there is no single solution that minimizes all the objectives simultaneously a set of solutions is required to be identified, where at least one of the objectives is better than the others. This set is called the *nondominated* (Pareto or efficient) set. In order to define the *nondominated* set, x is said to *dominate* y when

$$x < y \Leftrightarrow f_i(x) \leq f_i(y)(\forall\, i = 1 \ldots p) \wedge f_i(x) < f_i(y)(\forall\, i = 1 \ldots p) \tag{3}$$

One solution x is *nondominated* if there is no other feasible solution that dominates x. All the solutions in the *nondomi-*

nated set are candidates for selection. The general concept that lies behind the *nondominated* set is that along this set one objective is being traded off for another.

Multiobjective optimization methods should be able to generate the whole nondominated set. In reality this task is almost impossible, mainly due to the fact that it is very difficult to judge the completeness of the nondominated set, especially when a large number of objectives is involved.[27] In view of this, the analytical techniques utilized for the solution of multiobjective problems aim at finding an approximation of the nondominated set within the globally optimal domain. The most widely used of these techniques are the weighting and the constraint method.[30] The main idea behind the weighting method is that a set of weights is assigned to each objective function, and is then aggregated into a sum of objectives that is being optimized. On the other hand, the constraint method solves repeatedly a number of single objective optimization problems. One objective is optimized each time, whereas the rest of the objectives are turned into inequality constraints. There are conflicting reports in published literature with regards to the ability of the weighting method to target the globally optimal Pareto front. For example, Collete and Siarry[31] report that the weighting method presents difficulties in identifying all the solutions in the Pareto front and propose a combination of the weighting and the constraint method (that is, Tchebychev functions) in order to tackle these difficulties. In contrast Jaszkiewicz[32] finds that the weighting method is much more effective than the Tchebychev functions. Fu and Diwekar[27] report that the constraint method offers a more efficient way of exploring the Pareto front, and theoretically it can find all the solutions included in the Pareto front, but is computationally more expensive than the weighting method. Furthermore, there is usually no prior knowledge of the constraints utilized in the constraint method and sometimes no feasible solutions can be found for certain combinations of the constraints. Considering the above arguments, the weighting method is chosen in order to explore the Pareto front.

Weighting method

The idea behind the weighting method is to approximate the nondominated set through the identification of optimal points along the nondominated surface. Each objective function is associated with a weighting coefficient w and the weighted sum of the objectives is minimized. Therefore, the original problem defined by Eqs 1and 2 is now transformed as follows

$$\text{Minimize} \qquad \sum_{i=1}^{p} w_i f_i(x) \tag{4}$$

Subject to

$$x \in D \tag{5}$$

With

$$\sum_{i=1}^{p} w_i = 1 \tag{6}$$

According to Karush-Kuhn-Tucker (KKT) conditions, provided that all the weights are greater than zero, the optimal solution of the weighted problem (Eqs. 4 and 5) is a nondominated solution of the original problem (Eqs. 1and 2).[30] If the values of the weights are kept constant then a constant search direction will be achieved. The aim of MOO is to find all the possible tradeoffs among the objective functions. Therefore, variable weights are required in order to achieve multiple search directions and effectively explore the search space.[33] For a multiple objective problem with p objectives the weighted values can randomly be assigned as follows

$$w_i = \frac{rnd_i}{\sum_{j=1}^{p} rnd_j} \; (i = 1 \ldots p) \tag{7}$$

where rnd_i, rnd_j are nonnegative random real numbers in the range (0,1).

Multiple objective Simulated Annealing

Given the successful application of Simulated Annealing (SA) for single-objective molecular design,[19,20] this algorithm is used for the multiobjective case. The acceptance probability[34] is transformed in order to handle the multiple objectives. The employed acceptance probability function for multiple objective problems known as *scalar linear rule*[35] has the following form

$$\Pr = \min\left\{1, \exp\left(\frac{-\Delta F}{T}\right)\right\}, \tag{8}$$

where ΔF is given by

$$\Delta F = \sum_{i=1}^{p} w_i \Delta f_i \tag{9}$$

and

$$\Delta f_i = f_i(y) - f_i(x) \tag{10}$$

Each one of the objective functions must be normalized so that the acceptance probability is affected only by the use of weights. A logical flow diagram of multiobjective SA is presented in Figure 2.

During the optimization search an archive of a given number of nondominated solutions representing the dynamic Pareto optimality surface is recorded and updated. Therefore, the following archival rules can be implemented:[36]

- If the new solution dominates existing members of the archive, then they are removed and the new solution is added.
- If the new solution is dominated by any member of the archive, then it is not stored.
- If the new solution neither dominates nor is dominated by any member of the archive then:
  - If the archive is not full then the solution is stored.
  - If the archive is full then the new solution replaces the

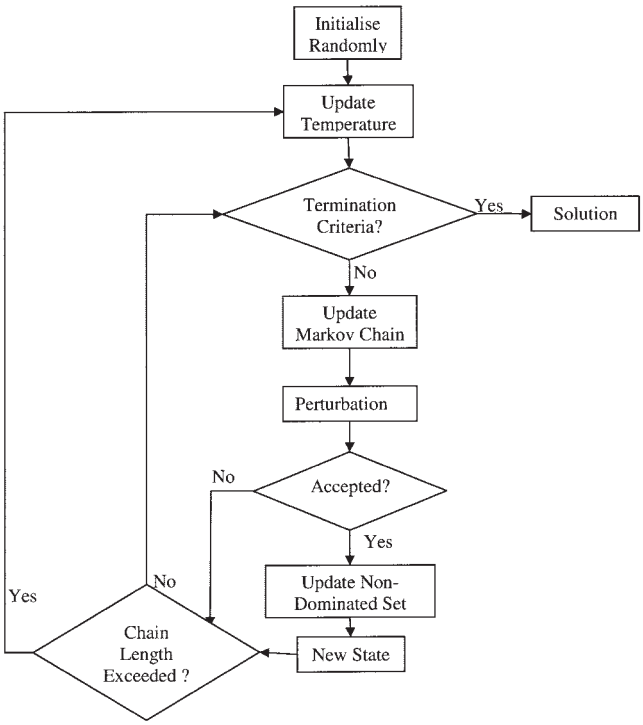

Figure 2. Flowchart for multiobjective SA.

last solution in the archive. Alternatively the most similar solution in the archive can be replaced, employing the difference in the objective function to measure the similarity.

Synthesis background and objectives

The solvent properties that need to be taken into account in the multiple objective functions employed by the proposed framework can be classified into three categories:[20]

1. Primary properties, such as critical properties, boiling-point temperatures, and so on, that are predicted using the group contribution methods proposed by Joback,[37] Joback and Reid,[38] Constantinou and Gani.[39]

2. Secondary properties, such as vapor pressure, heat of vaporisation etc., that are predicted using analytical expressions found in Reid et al.[40]

3. Mixture properties, such as activity coefficients involved in phase equilibrium, that are predicted using UNIFAC.

Process variables, such as stream flow rates, operating conditions and energy demands can be determined from process simulation models employing these properties.

Solvent performance indicators

A number of different mixture properties can be formulated as indicators of solvent-based separation process performance. Such indicators are calculated as a function of the activity coefficients at infinite dilution (γ^{∞}), and the molecular weights of the molecules (MW) [20]

$$\text{Relative volatility } a_{B,A} = \frac{\gamma^{\infty}_{B,S}}{\gamma^{\infty}_{A,S}} \cdot \frac{P^{B}_{vp}}{P^{A}_{vp}} \tag{11}$$

$$\text{Solvent selectivity } s_S = \frac{\gamma^{\infty}_{B,S}}{\gamma^{\infty}_{A,S}} \cdot \frac{MW_A}{MW_B} \tag{12}$$

$$\text{Solute distribution coefficient } M = \frac{\gamma^{\infty}_{A,B}}{\gamma^{\infty}_{A,S}} \cdot \frac{MW_B}{MW_S} \tag{13}$$

$$\text{Solvent losses in raffinate } S_l = \frac{1}{\gamma^{\infty}_{S,B}} \cdot \frac{MW_S}{MW_B} \tag{14}$$

$$\text{Solute solubility in solvent } S_b = \frac{1}{\gamma^{\infty}_{A,S}} \cdot \frac{MW_A}{MW_S} \tag{15}$$

where P^{B}_{vp},P^{A}_{vp} are the vapor pressures of the raffinate B and the extract A, respectively, and S denotes the solvent. The activity coefficients γ^{∞} are calculated at infinite dilution. Combinations of the above process performance indicators are generally used to formulate CAMD objective functions for separation solvent design.

The environmental impact can be measured by estimation of the toxicity (C_S) of each compound i using the group contribution technique of Gao et al.[41]

$$\text{Toxicity } (C_S) = -\log(LC50_i) = \sum_j n_{ij} a_j \tag{16}$$

where $LC50_i$ is the lethal concentration of component i (mol/L) causing 50% mortality in fathead minnow, and a_j is the contribution of group j.

Considering the molecular vector M_k that is adopted for the representation of the solvent molecules,[19] according to Eq. 4 the objective function can be formulated as

$$\text{Optimize } F(M_k) = \sum_{i=1}^{p} w_i f_i(M_k) + \sum_{i=1}^{m} Pen_i h_i(M_k) \tag{17}$$

where $f_i(M_k)$ are the synthesis objectives, $h_i(M_K)$ are the design constraints and Pen_i are the penalties associated with any design constraints.

Properties as synthesis objectives

The indicators of Eqs 11 through 15 reflect on desired solvent properties that would lead to good process performance. For instance, a desirable liquid-liquid extraction solvent would result in low solvent losses, be highly selective, have a low-distribution coefficient and a low-solute solubility. As there are always trade-offs between the different objectives, a compromise solvent is generally chosen. It is important for the user to understand these trade-offs, so as to appreciate the extents of the compromises that need to be made in solvent selection. Analysis of the Pareto optimal set determined by multiobjective optimization allows to capture the trade-offs between the different synthesis objectives.

The solvent performance indicators described in Eqs. 11 through 15 are functions of ratios of activity coefficients and molecular weights. This allows to trade off individual physical property objectives in a single performance measure, which is

Table 1. Sensitivity Analysis of Solvent Properties

Objectives	Solvent Dependent Properties			
	$\gamma^{\infty}_{S,B}$	$\gamma^{\infty}_{B,S}$	$\gamma^{\infty}_{A,S}$	MW_S
max S_S	—	max	min	—
max M	—	—	min	min
min S_l	max	—	—	min
max S_b	—	—	min	min

desirable from the point of view of single objective optimization in that a number of objectives can be captured in a single objective function. In contrast, the multiobjective optimization case considers each performance index as a separate design objective and would determine the trade-offs among the different indices. It is difficult for the user to comprehend the physical implications of the trade-offs between such performance indices, which again are weighted objectives of physical properties. This is because process design is based on physical properties directly rather than on solvent performance indicators. It would therefore be desirable to use the different physical properties directly as CAMD objectives. The translation of the performance indices into physical property objectives is discussed below.

The solvent performance indices employ the molecular weights of raffinate and extract in their definitions, which are independent of the particular solvent for a certain mixture. Thus, these (ratios of) molecular weights are parameters that have no effect on the outcome of the optimization. It is the solvent-specific physical properties ($\gamma^{\infty}_{B,S}$, $\gamma^{\infty}_{A,S}$, $\gamma^{\infty}_{S,B}$ and MW), which are important for CAMD and reflected in the four performance indices of Eqs. 12 through 15. Table 1 shows the relations of these performance indices to the physical property objectives. For instance, in order to maximize selectivity, $\gamma^{\infty}_{B,S}$ must be maximized and $\gamma^{\infty}_{A,S}$ must be minimized. Translation of performance indices into physical property objectives does not cause any conflicts, that is, multiobjective optimization-based CAMD employing the physical property objectives should identify the same Pareto front that would be obtained from optimization with performance index objectives. This means that multiobjective optimization-based CAMD allows the direct employment of physical properties, thus, creating a generic set of objective functions for which the Pareto optimal front retains all the important design information. Information on the selection of solvent performance indicators, which define the choice of property objectives, for different types of solvent-based separation processes is given by Marcoulaki and Kokossis.[20]

Illustrative examples

A number of studies are employed to present justification for decomposing the solvent-process systems design problem using a multiobjective CAMD approach, and to highlight the advantages of the approach over CAMD based on single-objective optimization. In particular, the applications intend to demonstrate the following points:

- Molecules obtained using the single objective optimization (SOO) synthesis approach as well as using simultaneous solvent-process design are included in the Pareto optimal front.
- The proposed MOO approach can sufficiently handle increasing numbers of objective functions.
- The use of solvent performance measures and property-based objectives leads to similar solvent design results.
- Molecules obtained using the SOO approach could potentially lead to suboptimal process performance.

The above points are demonstrated with seven studies spread over two examples that investigate the design of solvents for the separation of two mixtures. The first example, previously solved by Marcoulaki and Kokossis[20] using a SOO approach, involves the separation of the n-butanol-water system using liquid-liquid extraction. Case A compares the results obtained using the MOO approach with results previously proposed by Marcoulaki and Kokossis[20] for the same type of solvent performance indicators. The results obtained from the SOO and MOO approaches are compared again in cases B and C. The solvent performance indicators employed in the MOO formulation for case B are the properties shown in Table 1, whereas the environmental impact index is added in case C. Finally, the results obtained using the MOO approach are compared for cases A and B using statistical analysis techniques, and the effect of using increasing numbers of objectives in the MOO formulation is investigated.

The second example involves the separation of a mixture that consists of air and acetone. Marcoulaki and Kokossis[20] addressed this problem in order to demonstrate simultaneous solvent and process design. Case D presents the molecules obtained using the MOO objectives and compares these with

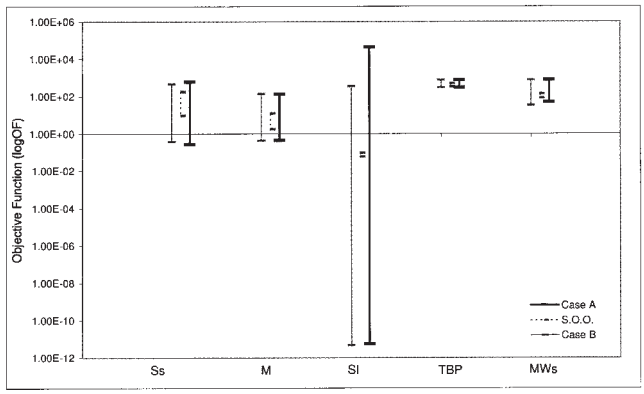

Figure 3. Case A vs. case B and SOO framework range of properties.

Table 2. Designs Obtained using MOO Solvent Performance Objectives for Case A

ID	Molecule	S_S	M	S_l	T_{BP}	MW_S
A1	$CH_2{=}C{-}(CH{=}CH_2)_2$	9.43	2.00	0.060	340	80
A2	$CH_2{=}CH{-}C(CH_3)_2{-}CH{=}CH_2$	14.21	1.49	0.021	353	96
A3	$CH_2{=}C(CH_3){-}CH(CH_3){-}CH{=}CH_2$	15.12	1.46	0.023	362	96
A4	$CH_3{-}CH_2{-}(CH_3)C(CH_2{-}CN){-}CH{=}CH_2$	71.53	11.42	0.096	467	123
A5	$CH_3{-}(CH_2)_4{-}C(CH_3)_2{-}CH_2{-}CN$	92.06	10.63	0.039	492	137
A6	$CH_3{-}CH_2{-}CH{=}CH{-}CH_2{-}CH(CH_3){-}CH_2{-}CN$	179.50	11.01	0.017	503	153

Table 3. Designs Obtained using MOO Property Objectives for Case B

ID	Molecule	S_S	M	S_l	T_{BP}	MW_S
B1	CH_3—$(CH_2)_3$—CH═C(CH_3)—CH_2—CN	103.20	11.38	0.049	497	137
B2	CH_3—$(CH_2)_3$—CH(CH_3)—CH(CH_3)—CH_2—CN	180.90	10.84	0.015	503	153
B3	CH_3—CH_2—(CH_3)C(CH_2—CN)—CH═CH_2	71.53	11.42	0.096	467	123
B4	CH_2═CH—CH(CH_3)—CH═CH_2	13.90	1.77	0.054	331	82
B5	CH_2═C(CH═CH_2)—CH═CH_2	9.43	2.00	0.060	340	80

results presented by Marcoulaki and Kokossis.[20] Cases E and F show that molecules synthesized using simultaneous solvent-gas-absorption design and solvent-gas-absorption-distillation design, respectively are captured by the Pareto optimal front. Case G shows that molecules designed using the MOO approach outperform molecules obtained using SOO in terms of process performance.

Simulated annealing is used with a Markov chain length of 100. The problems use a maximum of 20 UNIFAC groups in each state. Results are reported for a sample of 50 stochastic runs per case.

Example 1: Design of solvents for the separation of an n-Butanol-Water mixture using liquid-liquid extraction

Case A: MOO using solvent performance objectives. Cockrem and coworkers[42] propose that the most important performance objectives for the design of liquid-liquid extraction solvents are solute-distribution coefficient, solvent selectivity, solvent losses in the raffinate and boiling-point temperature. Molecular weight is also important as it accounts for the size of the molecule. The design objectives are the following:

$$\max S_S, M \tag{18}$$

$$\min S_l, MW_S \tag{19}$$

$$T_{\max} > T_{BP}^S > T_{BP}^A + \Delta T_{\min} \tag{20}$$

$$T_{\min} < T_{BP}^S < T_{BP}^A - \Delta T_{\min} \tag{21}$$

The boiling point temperature is addressed using a penalty function in order to be kept within the specified limits. Additional objectives can be employed, such as energy consumption in the regeneration of the solvent, or constraints that concern the structure of the molecules, if this kind of information is available.

In this first case Eqs.18 – 21 are considered as objectives in a MOO framework. Table 2 shows some of the designs included in the results. Molecules A1 and A4 are identical to molecules S1 and S8 proposed by Marcoulaki and Kokossis.[20] The remaining molecules are structurally similar to the ones proposed by Marcoulaki and Kokossis,[20] whereas they show reduced solvent losses and increased selectivities. Figure 3 shows that the property ranges of the solutions obtained using MOO include those of the solutions found previously by Markoulaki and Kokossis[20] using SOO.

Case B: MOO Using Property Objectives. Solvent performance indicators described by Eqs 11 – 15 involve certain weights that are only dependent on the mixture that is being separated, as previously explained. Retaining only the properties that are dependent on the solvent leads to a more general form of design objectives, and the optimization problem can be formulated as follows:

$$\max \gamma_{B,S}^{\infty}, \gamma_{S,B}^{\infty} \tag{22}$$

$$\min \gamma_{A,S}^{\infty}, MW_S \tag{23}$$

$$T_{\min} < T_{BP}^S < T_{BP}^A - \Delta T_{\min} \tag{24}$$

$$T_{\max} > T_{BP}^S > T_{BP}^A + \Delta T_{\min} \tag{25}$$

A sample of designs resulting from this formulation is shown in Table 3. Despite the generality of this formulation the results include molecules that are identical (B3, B4, B5) to the ones proposed by Marcoulaki and Kokossis[20] or molecules with similar properties and structures. The range of the results for both cases A and B as well as for the SOO approach is shown in Figure 3.

Case C: Environmentally Friendly Solvents Initially the problem of the design of environmentally friendly solvents for liquid-liquid extraction processes is addressed as a single objective optimization problem in order to create a reference dataset of designs with which the MOO results will be compared. The synthesis objectives remain the same as the ones proposed by Marcoulaki and Kokossis.[20] A constraint is added that accounts for toxicity in the designed molecules in the form of Eq. 16. In cases where there were no available group contribution data for the toxicity of certain groups in Gao et al.[41] the missing data were taken from Martin and Douglas.[43] For an increasing level of toxicity up to 5 the best results are presented in Table 4. Apparently, the results of the SOO

Table 4. Designs Obtained using SOO for Case C

ID	Molecule	C_S[a]	S_S	M	S_l	T_{BP}	MW_S
CS1	CH_2═CH—CH_2—(CH_2COO)$_2$—CH_2—CH═CH_2	1.03	7.24	1.54	0.045	511	198
CS2	CH_2═CH—CH_2—CH_2COO—CH_2—CH═CH_2	1.43	10.13	1.85	0.059	445	140
CS3	CH_2═CH—CH_2COO—CH═CH—CH═CH_2	1.51	7.32	1.63	0.054	445	138
CS4	CH_3—CH_2—CH_2—CH_2COO—CH═CH_2	1.85	11.26	1.92	0.068	445	140
CS5	CH_2═CH—(CH_2)$_4$—CH_2—CN	2.72	71.77	11.13	0.087	481	123

[a] Where $C_S = -\log(LC50_i)$

Table 5. Designs Obtained using MOO for Case C

ID	Molecule	C_S	S_S	M	S_l	T_{BP}	MW_S
CM1	CH_3—O—$(CH_2COO)_2$—CH═CH_2	0.09	1.49	1.31	3.860	487	174
CM2	CH_3COO—$(CH_2COO)_2$—CH_2—CH_2—O—CH_3	0.04	2.19	1.33	2.645	537	234
CM3	CH_3—$(CH_2COO)_3$—CH═CH—CH_2═CH	0.76	5.13	1.27	0.078	549	242
CM4	CH_2═CH—$(CH_2COO)_2$—CH═CH—CH_2═CH	1.11	5.46	1.32	0.037	511	196
CM5	CH_2═CH—CH_2—CH_2COO—CH_2—CH═CH_2	1.43	10.12	1.86	0.059	445	140
CM6	CH_2═CH—CH_2—CH_2—CH═CH_2	1.83	13.89	1.77	0.054	345	82

approach do not reach high environmentally benign limits nor do they demonstrate variability in their properties.

On the other hand, the MOO problem is formulated using Eqs 18 – 21, with *Cs* added in Eq. 19. The use of MOO leads to solutions with a wide range of structures and properties. Table 5 shows a sample of the variety of designs and trends that can be found in the results. Some of them involve very low toxicity, due to the intense presence of carboxyl or ether oxygen groups, but at the expense of low selectivity or high solvent losses. Others combine low toxicity with desirable values in the rest of the properties. The variability of the results obtained using MOO becomes apparent in Figure 4 that compares the range of the results obtained from the MOO and the SOO approach.

Comparative analysis and numerical results for cases A, B and C

Figures 5a, 5b, 5c show comparisons of samples from the set of solutions obtained for cases A, B, and C using the MOO approach with the designs proposed by Marcoulaki and Kokossis[20] denoted by S_i. The samples show that there are solutions with significantly improved selectivities and solvent losses whereas disimprovements in the other three objectives are very moderate. Figure 3 shows the full range of the solutions proposed in cases A and B in comparison with the range of the solutions reported by Marcoulaki and Kokossis.[20] Clearly, the ranges in approaches A and B are very similar and include all the solutions proposed by Marcoulaki and Kokossis.[20] On the other hand the ranges shown in Figure 4 for case C indicate a different overall set of molecules which should be expected as the existence of *Cs* introduces a new synthesis drive into the problem.

The existence of trends among the properties of the molecules becomes evident already from Table 2. These trends can be verified for the total range of the Pareto front using the correlation coefficient.[44] The correlation coefficient is widely used in molecular data analysis and shows the degree to which variables are related. In order to avoid making approximate assumptions regarding the normality of the distribution of the obtained data nonparametric statistical analysis is employed. Figure 6 shows the Spearman's rank correlation coefficient[45] for different properties in case A. A positive correlation coefficient indicates that as the rank of one property increases the rank of the other property increases as well. On the other hand, a negative correlation coefficient indicates that when the rank of one property increases the rank of the other property decreases. The use of the correlation coefficient also shows how strong or weak the correlations are. For example, Figure 6 shows that the trend between *Ss* and MW_S or is very weak but the trend between S_l and MW_S is very strong. It should be noted here that the results are verified by the conclusions of Cockrem

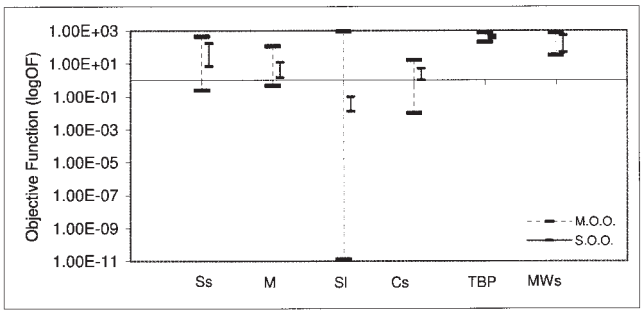

Figure 4. MOO vs. SOO range of properties for case C.

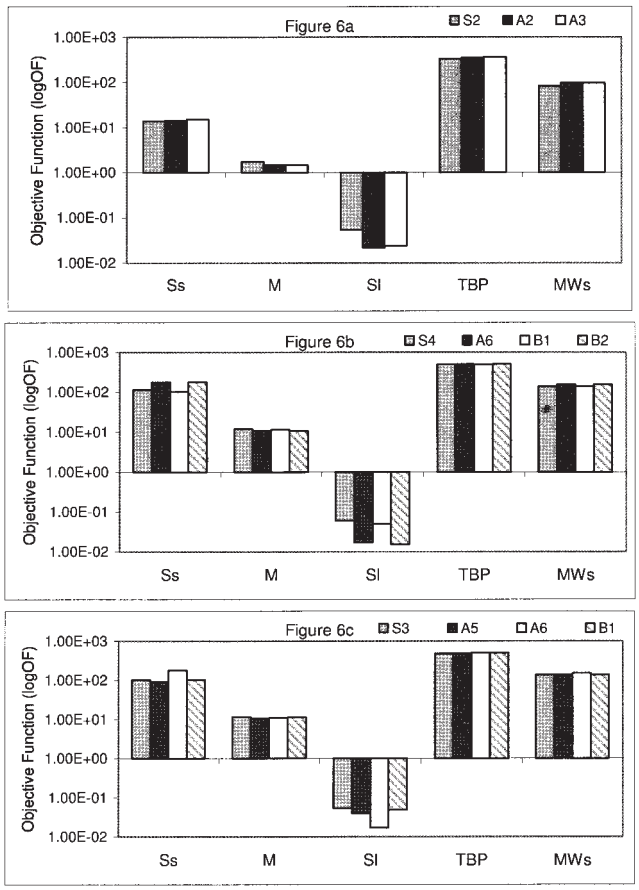

Figure 5. (a) Comparison of designs obtained by Marcoulaki and Kokossis[19] with designs obtained in case A; (b, c) comparison of designs obtained by Marcoulaki and Kokossis[19] with designs obtained in cases A and B.

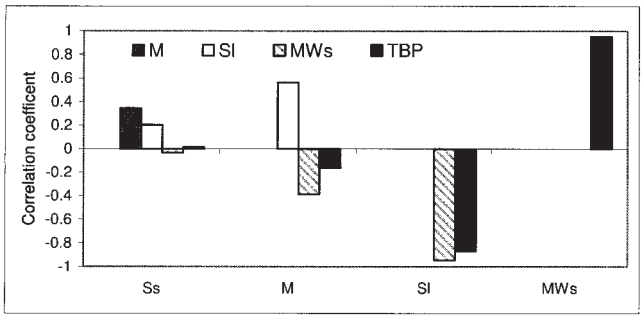

Figure 6. Correlation coefficient for properties of designed molecules in case A.

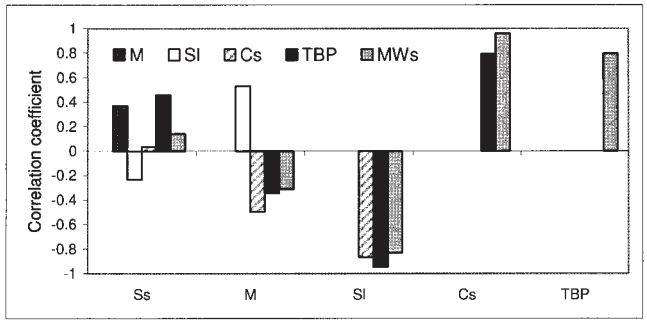

Figure 8. Correlation coefficient for properties of designed molecules in case C

et al. [42]. Figure 6 shows that a negative correlation exists between MW_S and M, and a negative correlation exists as well between MW_S and S_l. Cockrem et al. [42] find the same trends among the two pairs of properties. Using the correlation coefficient for case B, the results shown in Figure 7 indicate similar overall trends to the trends in case A. In case C the correlation coefficient shown in Figure 8 indicates different trends between the properties than the ones derived in cases A and B.

A detailed overall quantitative understanding of the relations between the decision variables can be gained if a data analysis tool is used. Cases A and B are of particular interest since there is evidence that the overall set of the designed molecules is quite similar. The analysis of the relationship among the five different objectives can be done using the multivariate regression analysis method in MATHCAD (www.mathcad.com). In this case a reasonable choice is to explore the relationship between S_S, and the other four objectives. The obtained nonlinear model is described by Eq. 26

$$S_s = \sum_{i=1}^{15} (coeffs_i \cdot M^{I_{i,1}} \cdot S_l^{I_{i,2}} \cdot MW_s^{I_{i,3}} \cdot T_{BP}^{I_{i,4}}) \qquad (26)$$

In Table 6 each element $I_{i,j}$ represents the order of the corresponding variable. The general model correlates the design objectives in a rather similar manner with differences only in the correlation coefficients. The sample multiple coefficient of determination R^2 indicates a reasonable fitting of the data into the model in both cases. Using the nonparametric test of two independent samples in SPSS (www.spss.com) it is possible to compare the two sets of coefficients for cases A and B and investigate whether their values differ significantly. In a level of significance of 5% there is a 57.6% possibility that the observed differences between the two sets of coefficients are due to random sampling error. This result indicates that the results follow the same distribution and supports the viewpoint that the models, and, thus, the designs obtained in cases A and B, are similar. Similar observations have been made in all other cases studied using this approach. Apart from property-property correlations, property — structure correlations can be identified from the results of the MOO approach. The results of the study of structure-property relations will be reported in a future publication.

By introducing an increasing number of objective functions into the MOO framework, its performance with respect to computational effort is investigated. The average CPU time and number of function evaluations per run are shown in Table 7. The computational time and the function evaluations remain at low levels even in the case of 5 objectives.

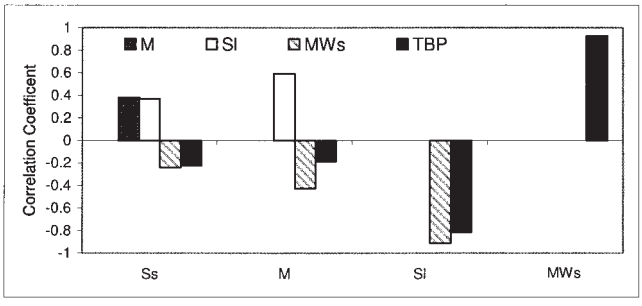

Figure 7. Correlation coefficient for properties of designed molecules in case B.

Example 2: Design of solvents for the separation of an Air-Acetone mixture using gas-absorption

Case D: Design objectives for MOO. Let us assume that some important properties for the design of solvents suitable for gas-absorption are:

- Solute solubility in solvent as it determines the amount of the input solvent flow rate.

Table 6. Order of Polynomial and Coefficients for Cases A and B

	Order I of Polynomial				Coefficients for Case A R^2: 90.5%	Coefficients for Case B R^2: 90.2%
	1	2	3	4		
1	1	0	0	1	−0.9	−0.611
2	0	1	0	1	−0.451	11.587
3	0	0	1	1	−0.024	−0.041
4	0	0	0	2	0.0121	0.0137
5	0	0	0	1	−5.332	−2.529
6	1	0	1	0	0.4663	0.5824
7	0	1	1	0	0.5864	−8.127
8	0	0	2	0	0.0092	0.0324
9	0	0	1	0	7.8568	3.4193
10	1	1	0	0	−61.66	−82.19
11	0	2	0	0	89.672	583.07
12	0	1	0	0	134.14	−4497
13	0	0	0	0	0.0013	−0.003
14	1	0	0	0	388.5	209.77
15	2	0	0	0	8.3895	0.2251

Table 7. Numerical Results for Increasing Number of Objectives

Number of Objectives	CPU Time (sec)	Function Evaluations
3	75	144193
4	136	251879
5	158	291648

- To facilitate separation in solvent regeneration the boiling points of the solvents should be significantly higher or lower than those of the nonsolvent molecules that are to be absorbed.[20] Furthermore, low solvent boiling points imply high solvent loss, whereas high solvent boiling points imply high relative volatility, low chance of azeotrope formation and high energy demand in regeneration. Due to the low boiling point of acetone, only solvent molecules with boiling points greater than those of the nonsolvent molecules were developed in this study.
- Low volatility (represented by the vapor pressure of the solvent) as it reduces the loss of solvent and facilitates solvent regeneration.
- Heat of vaporisation as it determines the energy demand in regeneration.
- Low-molecular-weight implies smaller molecules that are simpler to design.

The design objectives are

$$\max S_b \tag{27}$$

$$\min H_v, MW_S, P_{vp} \tag{28}$$

$$T_{BP}^S > T_{BP}^A + \Delta T_{\min} \tag{29}$$

$$T_{BP}^S > T_{BP}^{Water} \tag{30}$$

$$T_{BP}^S < T_{\max} \tag{31}$$

This separation is carried out using water which must be replaced by a more efficient solvent. The properties are calculated at normal conditions with the exception of the vapor pressure of the solvent (P_{vp}). A realistic prediction of the vapor pressure requires bubble point calculations at the bottom of the column. Due to these calculations the equilibrium liquid concentration X^* for a gas featuring Y in the vapor phase is available. It is, therefore, preferable to use X^* in Eq. 27, instead of representing solubility with Eq. 15. The results of the MOO framework are presented in Table 8. They involve many simple molecular structures that can be easily synthesized. Marcoulaki and Kokossis[20] combined the boiling point constraint with solvent flow rate and operating temperature as process economic criteria in their simultaneous process-solvent design approach. It is worth noting that the majority of the designs obtained by that approach are included in Table 8. This is a first notion of proof that the designs obtained using MOO include molecular structures that are designed when process feedback is present.

Case E: Simultaneous Gas-Absorption-Solvent Synthesis

In this case a gas-absorption process is designed in course of solvent synthesis by adding two constraints regarding the formation of binary azeotropes and the solvent molecular weight on the representation proposed by Marcoulaki and Kokossis.[20] The design decisions in the optimization involve the solvent flow rate and the structural solvent decisions. The representation of Marcoulaki and Kokossis[20] includes the following predefined process parameters:

- The number of theoretical stages N_{ab} in the column.
- The operating pressure P in the column.
- The inlet y_{in} and outlet y_{out} concentration of acetone in the gas phase.

Additional constraints include bounds on the solvent mass flow rate and the operating temperature T_{ab} at the bottom of the column

$$S_{ab}^{FL} < S_{ab,\max}^{FL} \tag{32}$$

$$T_{ab} > T_{ab,\min} \tag{33}$$

The bounds with respect to the solvent involve molecular weight MW_S, formation of binary azeotropes and boiling point T_{BP}

$$MW_S < MW_{S,\max} \tag{34}$$

$$(P_{vp}^i - \gamma_{ii,i}^{\infty} \cdot P_{vp}^{ii}) \cdot (P_{vp}^{ii} - \gamma_{i,ii}^{\infty} \cdot P_{vp}^i) < 0 \tag{35}$$

where P_{vp} and γ^{∞} are the vapor pressure and the infinite dilution activity coefficient for components *1* and *2*, respectively. Constraint (Eq. 35), accounting for the formation of binary azeo-

Table 8. Designs Obtained using the MOO Framework for Gas-Absorption

ID	Molecule	X^*	H_v	MW_S	T_{BP}	P_{vp}
C1	$CH_2{=}CH{-}CH_2{-}COOH$	0.34	48.9	86	438	$0.33 * 10^{-5}$
C2	$CH{\equiv}C{-}CH(CH_3){-}CH_2{-}NO_2$	0.24	41.0	113	463	$0.27 * 10^{-5}$
C3	$CH{\equiv}C{-}CH_2{-}CH_2{-}CH_2{-}NO_2$	0.16	42.2	113	470	$0.25 * 10^{-5}$
C4	$CH_3{-}CH_2{-}CH_2{-}CN$	0.05	33.6	69	394	$0.17 * 10^{-1}$
C5	$CH_3{-}CH{=}CH{-}C(=O){-}CH_3$	0.04	32.8	84	376	$0.35 * 10^{-1}$
C6	$CH_3{-}COOH$	0.05	46.6	60	389	$0.13 * 10^{-1}$
C7	$FCH_2{-}O{-}CH_2{-}CH_2{-}Cl$	0.11	31.7	94	382	$0.93 * 10^{-2}$
C8	$CH_2{=}CH{-}(Cl)C(-O{-}CH_3){-}CH_2{-}CN$	0.10	43.2	146	495	$0.16 * 10^{-5}$
C9	$CH_2{=}CH{-}CH(-O{-}CH_3){-}Cl$	0.21	30.8	106	389	$0.12 * 10^{-2}$
C10	$FCH_2{-}O{-}CH_2{-}CH(-O{-}CH_2F){-}Cl$	0.25	36.3	123	444	$0.28 * 10^{-5}$
C11	$CH_3{-}CH(HC{=}O){-}Cl$	0.06	22.1	93	380	$0.19 * 10^{-1}$
C12	$Cl{-}CH(HC{=}O){-}C(-O{-}CH_3){=}CH_2$	0.18	38.0	129	460	$0.19 * 10^{-5}$
C13	$CH{\equiv}C{-}CH_2{-}NO_2$	0.06	36.9	85	428	$0.79 * 10^{-3}$

Table 9. Designs Obtained for Sim96qultaneous Solvent Design with Gas Absorption

ID	Molecule	S_{ab}^{FL} (ton/hr)	T_{ab} (K)	MW_S	T_{BP}
D1	CH_2═CH—CH_2—COOH	21.9	297	86	438
D2	CH_2═CH—(Cl)C(CH_3)—CH_2—CN	40.9	297	130	469
D3	CH≡C—$CH(CH_3)CH_2$—NO_2	41.9	309	113	463
D4	FCH_2—O—CH_2—CH(—O—CH_2F)—Cl	42.8	306	123	444
D5	CH_2═CH—CH(—O—CH_3)—Cl	43.6	295	106	390
D6	CH_3—COOH	107.3	321	60	389

tropes, has been proposed by Buxton, Livingston and Pistikopoulos.[14] The bounds for T_{BP} remain the same as in Eqs 29 – 31. Apart from the fact that the use of constraint (Eq. 35) adds different search directions to the optimization problem, it is required in order to account for molecules that facilitate the separation during the regeneration of the solvent. The synthesis objective is to minimize the solvent mass flow rate (S_{ab}^{FL}) in the process, and the model is solved simultaneously with the solvent design model. The set of results is generated by increasing the bound in MW_S after each run. The results presented in Table 9 are ranked, based on increasing S_{ab}^{FL} and are very similar to those proposed by Marcoulaki and Kokossis.[20]

All the designs have been identified in Table 8 as part of the Pareto front, apart from D2, which is, however, structurally similar to C8. Furthermore, Figure 9 shows that all the solutions found in case E lay within the range of the solutions obtained using the MOO approach. Even though the constraints in this case are different to the MOO constraints, the MOO method is highly inclusive in optimal designs. This shows that the MOO methodology can capture solutions that are generated when process performance feedback is present.

Case F: Simultaneous Gas-Absorption-Distillation-Solvent Synthesis In this case a distillation column is added to the existing gas absorption process in order to recover the used solvent. Figure 10 shows a representation of the absorption-distillation process. The equations and assumptions for the design of the distillation column using a shortcut method are reported in Appendix A. Additional predefined process parameters for the distillation column are:

- The mole fraction x_b of acetone at the bottom of the column.
- The mole fraction x_d of acetone at the top of the column.

Constraints for gas-absorption follow Eqs 32 – 35. Constraints for the distillation column include bounds on the number of stages N_{dst}, and on the reflux ratio R_{dst}

$$N_{dst} < N_{dst,\max} \tag{36}$$

$$R_{dst} > R_{dst,\min} \tag{37}$$

The synthesis objective is to minimize the energy E_{dst} required for the recovery of the solvent. The model is solved simultaneously with the solvent design model. The results are ranked based on increasing E_{dst}. Figure 9 shows that the results obtained using the MOO methodology are inclusive of the designs obtained when process performance feedback is accounted for at the solvent design stage. The optimum design E1 in Table 10 has been identified in Table 8 as part of the Pareto front. Designs E2 and E3 are structurally similar to designs C12 and C13, whereas the rest of the designs are highly suboptimal and violate the boiling point constraint. The earlier arguments show that the MOO methodology can capture the exact designs, as well as the synthesis trends that exist when process performance feedback is present at the solvent design stage.

Case G: Integration of solvent design results into gas-absorption-distillation synthesis The solvent molecules obtained at the MOO solvent design stage are introduced as discrete variables into the gas-absorption-distillation synthesis model. The process optimization variables are:

- The number of theoretical stages N_{ab} and N_{dst} in the absorption and distillation column, respectively.
- The outlet concentration of acetone y_{out} in the gas phase.

Constraints for the gas-absorption and the distillation column follow Eqs 32, 33 and 37. Furthermore, bounds are set for the operating temperature of the distillation column $T_{\rm dst}$, and the absorption factor AF

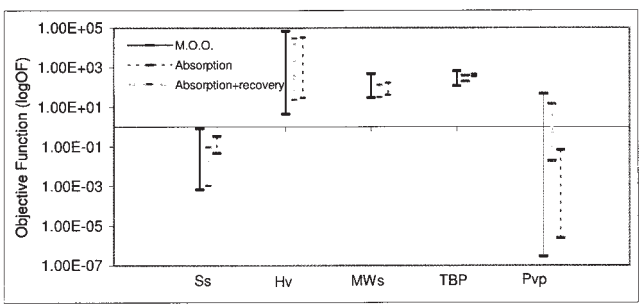

Figure 9. Ranges of properties for MOO vs. SOO simultaneous solvent-process design approach.

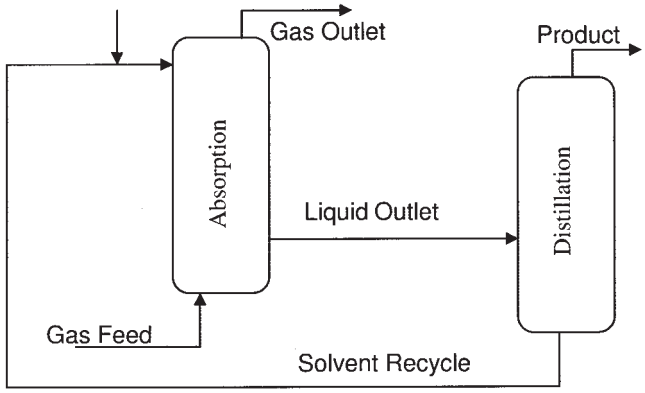

Figure 10. Integrated solvent-process design for separation of air-acetone mixture.

Table 10. Designs Obtained for Simultaneous Solvent Design with Gas Absorption and Recovery

ID	Molecule	E_{dst} (MJ/hr)	S_{ab}^{FL} (ton/hr)	T_{ab} (K)	N_{dst}	R_{dst}	MW_S	T_{BP}
E1	CH_3—CH(HC═O)—Cl	6052	135.8	317	11	1.03	93	381
E2	CH_2═C(HC═O)—O—CH_3	6285	295.3	330	8	1.06	86	392
E3	CH_3—CH_2—NO_2	7162	140.5	319	10	1.07	75	387
E4	HO—CH_2—CL	9701	214.0	325	10	1.63	66	372
E5	CH_3—CH_2—CN	12330	71.7	312	18	3.27	55	364
E6	CH_3—O—OH	21770	302.5	322	11	5.31	48	347
E7	CH_2═CH—OH	45440	87.97	311	31	1.33	44	329

$$T_{dst} > T_{dst,\min} \tag{38}$$

$$AF_{\min} < AF < AF_{\max} \tag{39}$$

The optimum molecule obtained in this case is shown in Table 11 in comparison with the molecule obtained in the SOO case without process feedback.

The SOO molecule was obtained by using X^* as the main objective function, whereas the rest of the properties were considered as additional constraints. In this respect X^* was allowed to reach a rather high limit and the molecule should represent the optimum solvent for an absorption-distillation process. However, the optimum molecule in terms of process cost has different properties as it lies in an area of the solvent synthesis solution space where X^* is low, and the vapor pressure is significantly lower than the SOO molecule. Figure 11 shows a correlation of the process cost (OF) of the molecules that belong to the Pareto front with the different physical properties.

The vapor pressure has a stronger effect on the process cost than the equilibrium concentration X^*. Thus, by using X^* as the objective function in the SOO solvent design problem, the search is dominated by solutions that obtain a high X^*, and the optimum solution in terms of process cost is lost. For the same reason using any other property as a single objective function would potentially lead to suboptimal results. With regards to the molecular weight Figure 11 shows that heavier molecules lead to reduced process cost. In the SOO case the molecular weight was included in the search as an additional constraint with an upper bound. In this respect using a low value as the upper bound would lead to suboptimal results.

Concluding Remarks

This work presented a new methodology for the design of molecules that can be utilised as solvents in various processes. It builds upon multiple objective optimization in order to generate a set of optimal solvent molecules that embeds highly inclusive designs and incorporates comprehensive design information regardless of process economics.

Overall the results obtained from the utilized examples show that multiple objective optimization in combination with CAMD can be a very useful tool for the design of solvents. Unlike previous approaches that focused on one final, optimum design, it has allowed the generation of a rich, highly inclusive set of designs. A comparison with solutions found previously for SOO indicates that they are included in the designs obtained using the MOO methodology. A variety of performance and environmental indices is used with success whereas the complexity of the design problem is not significantly increased by including additional solvent property objective functions. The obtained set of noninferior designs allows the hidden, complex, intermolecular trends to be revealed, thus providing with design insights through the identification of patterns. The presented methodology is able to operate at a conceptual level, where the solvent properties are expressed in a generic fashion avoiding possible biases toward certain processes during the solvent design. As the different synthesis drives are treated as independent objective functions, the important process information otherwise required at the solvent design stage in the form of process feedback is incorporated in the results. Marcoulaki and Kokossis[20] observe that when process performance feedback is present at the solvent design stage controls for molecular complexity are unnecessary. In the presented methodology controls and artificial constraints become redundant as the obtained information involves both simple and complex designs.

The study presented in this work introduces the MOO solvent synthesis concept. In this context the technology employed for the implementation of the proposed MOO framework is simple, but sufficient enough in order to lead to successful solvent design results. However, there are issues with regards to the identification of the Pareto front that are not addressed in this study. For example, choosing the weighting method in order to approximate the Pareto front leads to molecules that are similar, identical or better than molecules identified using SOO or process performance feedback in course of solvent design. However, there is no guarantee that the obtained set of Pareto points is evenly distributed into the solvent solution space. Furthermore, molecules that have not been identified might exist in the Pareto front. In this respect more sophisticated methods that address the effective identification of solutions in the Pareto front should be trialled. One such method, proposed by Fu and Diwekar,[27] utilizes optimization under uncertainty in order to efficiently capture Pareto optimal points.

Table 11. Comparison of SOO and MOO Solvent Results Based On Process Performance

Case	Molecules	X^*	H_v	MW_S	P_{vp}	T_{BP}	*Cost* (\$/yr)
M.O.O.	CH_3—O—CH(Br)CH(Br)Cl	0.073	41.1	252	$7.03 * 10^{-7}$	487	486010
S.O.O.	CH_2═C(—O—CH_3)—CH_2—Cl	0.855	32.4	106	$2.86 * 10^{-5}$	395	657729

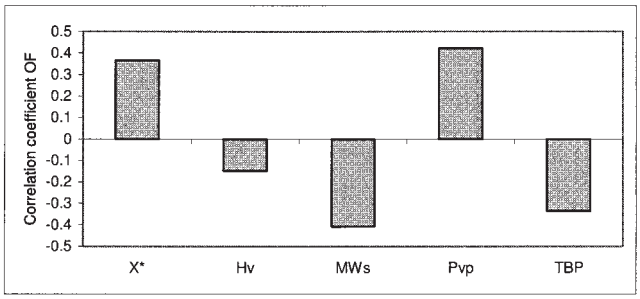

Figure 11. Correlation coefficient for cost (OF)– solvent properties in integrated solvent-process design.

Literature Cited

1. Brignole EA, Botini S, Gani R. A strategy for the design and selection of solvents for separation processes. *Fluid Phase Equilibria*. 1986;29: 125.
2. Joback KG, Stephanopoulos G. *Designing molecules possessing desired physical property values*. Proc. FOCAPD, CACHE Corp., Austin, TX; 1989;11:631.
3. Gani R, Nielsen B, Fredenslund, A. Group contribution approach to molecular design. *AIChE J*. 1991;37:1318.
4. Pretel EJ, Lopez PA, Botinni SB, Brignole EA. Computer aided molecular design of solvents for separation processes. *AIChE J*. 1994; 40:1349.
5. Harper PM, Gani R. A multi-step and multi-level approach for computer aided molecular design. *Comps and Chem Eng*. 2000;24:677.
6. Hostrup M, Harper PM., Gani R. Design of environmentally benign processes: Integration of solvent design and separation process synthesis. *Comps and Chem Eng*. 1999;23:1395.
7. Vaidyanathan R, El-Halwagi M. Computer-aided design of high performance polymers. *J of Elastomers and Plastics.* 1994;26:277-293.
8. Duvedi AP, Achenie LEK. Designing environmentally friendly refrigerants using mathematical programming. *Chem Eng Sci*. 1996;51: 3727-3739.
9. Churi N, Achenie LEK. Novel mathematical programming model for computer-aided molecular design. *Ind and Eng Chem Res*. 1996;35: 3788-3794.
10. Sahinidis VN, Tawarmalani M. Applications of global optimisation to process and molecular design. *Comps and Chem Eng*. 2000;24:2157-2169.
11. Odele O, Machietto S. Computer aided molecular design: A novel method for optimal solvent selection. *Fluid Phase Equilibria*. 1993; 82:337.
12. Ismail SR, Pistikopoulos EN, Papalexandri KP. Separation of non-ideal mixtures based on mass/heat exchange principles: The entrainer selection and sequencing problem. *Comp and Chem Eng.* 1997;21: S211.
13. Pistikopoulos EN, Stefanis SK. Optimal solvent design for environmental impact minimization. *Comp and Chem Eng.* 1998;22:717-733.
14. Buxton A, Livingston AG, Pistikopoulos EN. Optimal design of solvent blends for environmental impact minimization. *AIChE J*. 1999; 45:817-843.
15. Wang Y, Achenie LEK. Computer aided solvent design for extractive fermentation. *Fluid Phase Equilibria*. 2002; 201:1.
16. Giovanoglou A, Barlatier J, Adjiman CS, Pistikopoulos EN, Cordiner J. Optimal solvent design for batch separation based on economic performance. *AIChE J.* 2003;49:3095.
17. Venkatasubramanian V, Chan K, Caruthers JM. Computer Aided Molecular Design Using Genetic Algorithms. *Comp and Chem Eng.* 1994;18:833-844.
18. Van Dyk B, Nieuwoudt I. Design of solvents for extractive distillation. *Ind and Eng Chem Res.* 2000;39:1423-1429.
19. Marcoulaki EC, Kokossis AC. On the development of novel chemicals using a systematic optimization approach. Part II. Solvent Design. *Chem Eng Sci*. 2000;55:2547.
20. Marcoulaki EC, Kokossis AC. On the development of novel chemicals using a systematic synthesis approach. Part I. Optimization framework. *Chem Eng Sci*. 2000;55:2529.
21. Marcoulaki EC, Kokossis AC, Batzias FA. Novel chemicals for clean and efficient processes using stochastic optimization. *Comp and Chem Eng.* 2000;24:705-710.
22. Kim KJ, Diwekar UM. Integrated solvent selection and recycling for continuous processes. *Ind and Eng Chem Res*. 2002;41:4479.
23. Hamad AA, El-Halwagi MM. Simultaneous synthesis of mass separating agents and interception networks. *Transactions of IChemE (Part A).* 1998;76:376.
24. Wang Y, Achenie LEK. A hybrid optimization approach for solvent design. *Comp and Chem Eng*. 2002;26:1415.
25. Eden MR, Jorgensen SB, Gani R, El-Halwagi MM. A novel framework for simultaneous separation process and product design. *Chem Eng and Proc*. 2004;43:595.
26. Linke P, Kokossis AC. *Simultaneous Synthesis and Design of Novel Chemicals and Process Flowsheets*. In: Grievink J, van Schijndel K. Computer-Aided Chemical Engineering 10. Elsevier; 2002:115-120.
27. Fu Y, Diwekar U. An efficient sampling approach to multiobjective optimization. *Annals of Operation Res*. 2004;132:109-134.
28. Dantus MM, High KA. Evaluation of waste minimization alternatives under uncertainty: a multiobjective optimization approach. *Comp. and Chem Eng*. 1999;23:1493-1508.
29. Silva CM, Biscaia EC. Genetic algorithm development for multiobjective optimization of batch free-radical polymerization reactors. *Comp and Chem Eng.* 2003;27:1329-1344.
30. Diwekar U. *Introduction to Applied Optimization*. Boston: Kluwer Academic Publishers; 2003.
31. Collete Y, Siarry P. Three new metrics to measure the convergence of metaheuristics towards the Pareto frontier and the aesthetic of a set of solutions of biobjective optimization. *Comp and Operations Res*. 2005;32:773-792.
32. Jaszkiewicz A. Genetic local search for multi-objective combinatorial optimization. *Euro J of Operational Res*. 2002;137:50-71.
33. Murata T, Ishibuchi H, Tanaka H. Mutli-objective genetic algorithm and its applications to flowshop scheduling. *Comp and Ind Eng.* 1996;30:957-968.
34. Metropolis NA, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations for fast computing machines, *J of Chem Phys*. 1953;21:1087.
35. Kubotani H, Yoshimura K. Performance evaluation of acceptance probability functions for multi-objective SA. *Comp and Operation Res*. 2003;30:427-442.
36. Busacca GP, Marseguerra M, Zio E. Multiobjective optimization by genetic algorithms: application to safety systems. *Reliability Eng and System Safety*. 2001;72:59-74.
37. Joback KG. *Designing Molecules Possessing Desired Physical Property Values.* Cambridge, MA:MIT; 1989. Ph.D. thesis.
38. Joback KG, Reid RC. Estimation of pure component properties from group-contributions. *Chem Eng Communications*. 1987;57:233-243.
39. Constantinou L, Gani R. New group contribution method for estimating properties of pure compounds. *AIChE J*. 1994;40:1697-1710.
40. Reid RC, Prausnitz JM, Polling BE. *The Properties of Gases and Liquids.* 4th ed. New York:McGraw-Hill; 1987.
41. Gao C, Govind R, Tabak HH. Application of the group contribution method for predicting the toxicity of organic chemicals. *Environ Technol and Chem.* 1992;11:631-636.
42. Cockrem M, Flatt J, Lightfoot E. Solvent selection for extraction from dilute solution. *Separation Sci and Technol*. 1989;24:769-807.
43. Martin MT, Douglas MY. Prediction of the acute toxicity (96-h LC_{50}) of organic compounds to the fathead minnow (Pimephales promelas) using a group contribution method. *Chem Res in Toxicology*. 2001;14: 1378-1385.
44. Livingstone D. *Data Analysis for Chemists*. Oxford Science Publications: New York; 1995.
45. Mendenhall W. *Statistics for the engineering and computer sciences.* 2nd ed. San Fracisco: Dellen Pub, Co;1988.
46. Chang HY. Gilliland plot in one equation. *Hydrocarbon Processing*. 1981;60:146.
47. Stichlmair JG, Fair JR. *Distillation, Principles and Practice*. Wiley-VCH; 1998.

Appendix A

Design of distillation column

The solvent is recovered after the absorption using a distillation column that is modeled with shortcut methods. The embedded assumptions of the shortcut models involve constant molar flow rates and relative volatility throughout the column, the feed, and the reflux being saturated liquid, phase equilibrium modeled using UNIFAC and ideal equilibrium stages. Furthermore, the heat required in the reboiler is assumed to be equal to the heat to be removed in the condenser and the condenser is considered total. The mass fractions of the distillate and the bottom streams are fixed.

Minimum number of theoretical stages

The minimum number of theoretical stages is calculated using the Fenske equation

$$\frac{x_D}{1-x_D} = a^{N_{\min}} \cdot \left(\frac{x_B}{1-x_B}\right) \tag{A1}$$

where x_D and x_B are the mole fraction in the distillate and bottom, respectively, and a is the relative volatility for a binary mixture.

Minimum reflux ratio

The Underwood equation is used for the calculation of the minimum reflux ratio

$$\frac{ax_F}{a-\theta} + \frac{1-x_F}{1-\theta} = 1-q \tag{A2}$$

$$r_{\min} = \frac{ax_D}{a-\theta} + \frac{1-x_D}{1-\theta} - 1 \tag{A3}$$

where α is the relative volatility, q is the caloric factor and ϑ is the parameter connecting the two equations.

Number of theoretical stages

After having calculated the minimum number of theoretical stages N_{min}, the real number of theoretical stages N can then be calculated from the Gilliland diagram. The following correlation derived from the Gilliland diagram is used[46]:

$$Y = 1 - \exp\left[1.490 + 0.315 \cdot X - \frac{1.805}{X^{0.1}}\right]$$
$$X = \frac{r - r_{\min}}{r+1}$$
$$Y = \frac{N - N_{\min}}{N+1} \tag{A4}$$

Heat requirements

The quantity of heat that is removed in the condenser at the top of the column can be calculated using the Eq.[47]

$$Q_c = F \cdot H_{vap} \cdot (R+1) \cdot \frac{z_F - x_F}{x_D - x_B} \tag{A5}$$

where F is the flow rate of the feedstream, H_{vap} is the heat of vaporization of the less volatile compound, R is the reflux ratio, z_F, x_D and x_B are the mole fractions of the most volatile compound in the feed, the distillate and the bottom products, respectively. The required heat in the reboiler is considered to be equal to the heat removed in the condenser ($E_{dst} = Q_C$).

Manuscript received Dec. 12, 2004, and revision received Sept. 16, 2005.