Optimal Design of Solvent Blends for Environmental Impact Minimization

Andrew Buxton, Andrew G. Livingston, and Efstratios N. Pistikopoulos
Dept. of Chemical Engineering, Centre for Process Systems Engineering, Imperial College, London SW7 2BY, U.K.

A systematic procedure proposed here selects optimal solvent blends for nonreactive, multicomponent absorption processes accounting for plant-wide point source environmental interactions. This approach, based on the optimal design technique for pure solvents developed in our earlier work, involves the identification of all agent-based operations (such as gas absorption and liquid extraction) within the process of interest, the determination of a list of candidate solvent blends satisfying separation task operational and environmental constraints, and the selection of an optimal solvent blend candidate based on global plant-wide process and environmental constraints. The procedure features new property and operational feasibility constraints, a refined task-based problem formulation, and a new solution strategy. Two examples are presented to demonstrate the approach, and trade-offs between cost and environmental impact objectives are discussed.

Introduction

The systematic design of molecules with desired properties is a key objective in the process industries. Computer-aided molecular design (CAMD) (Gani et al., 1991) techniques are increasingly being used to identify promising candidates from among the enormous range of potential alternatives. Such techniques are intended to develop molecules *directly* according to property and performance specifications, most commonly by employing a reverse engineering approach of systematically combining sets of structural groups.

In recent years, significant advances have been made in the use of methods to predict physico-chemical and environmental properties for efficient molecular design (Fredenslund et al., 1975; Joback and Stephanopoulos, 1989; Mavrovouniotis, 1990; Gao et al., 1992; Constantinou et al., 1994). These methods have been employed in the computer-aided design of polymers (Venkatasubramanian et al., 1994; Vaidyanathan and El-Halwagi, 1996), extractants (Gani and Brignole, 1983; Macchietto et al., 1990; Naser and Fournier, 1991), solvents (Brignole et al., 1986; Pretel et al., 1990; Klein et al., 1992; Gani and Fredenslund, 1993; Odele and Macchietto, 1993; Pistikopoulos and Stefanis, 1998), solvents in reactive systems (Modi et al., 1996), and refrigerants (Duvedi and Achenie, 1996b; Churi and Achenie, 1996).

The solution algorithms employed include enumeration techniques (Derringer and Markham, 1985; Stephanopoulos and Townsend, 1986; Joback, 1989; Joback and Stephanopoulos, 1989, 1995), knowledge-based approaches (Brignole et al., 1986; Klein et al., 1992; Gani et al., 1991, 1993; Modi et al., 1996), graph reconstruction methods (Gordeeva et al., 1990; Kier et al., 1993), multistage strategies (Naser and Fournier, 1991; Gani and Fredenslund, 1993), genetic algorithms (Venkatasubramanian et al., 1994), artificial intelligence (Bollis et al., 1991), local MINLP optimization (Macchietto et al., 1990; Odele and Macchietto, 1993; Vaidyanathan and El-Halwagi, 1996; Duvedi and Achenie, 1996b; Pistikopoulos and Stefanis, 1998), mixed integer linear optimization for linear structure-property relations (Constantinou et al., 1996), and exact reformulations for specific nonlinear structure-property relations (Maranas, 1996) later extended to quantify the effect of uncertainty on optimal molecular design (Maranas, 1997).

Environmental objectives have been included both implicitly (Constantinou et al., 1994) and explicitly (Duvedi and Achenie, 1996b; Pistikopoulos and Stefanis, 1998).

Correspondence concerning this article should be addressed to E. N. Pistikopoulos.

Despite the plethora of material design techniques, few have been applied to the problem of optimal material *blend* design. However, as processing and environmental demands become ever more stringent, the identification of appropriate single molecules becomes ever more difficult. Material blend design provides access to a wider range and combination of properties than single materials can offer. Furthermore, material blends can be precisely tuned to meet the increasingly demanding specifications.

Vaidyanathan and El-Halwagi (1996) recently applied CAMD to polymer blend design. Joback (1989), Macchietto et al. (1990), Klein et al. (1992), Gani and Fredenslund (1993) and Dunn et al. (1997) considered *solvent* blend design, while Duvedi and Achenie (1996a) considered *refrigerant* mixture design. Joback proposed an interactive (nonautomated) approach to group-based solvent blend design, while Macchietto et al. put forward a UNIFAC-based technique for solvent blend design but applied it only with predetermined molecules. The remaining authors extracted their compounds from a database of complete compounds. Solvent blend design represents a key process engineering activity since solvents are used pervasively in the chemical industries to effect separations, store raw materials, quench reactions, purify products, clean process equipment, provide media for crystallization, and so on. In this article, we present an algorithmic, group-based, computer-aided molecular design approach for the optimal design of solvent blends for nonreactive, multicomponent absorption.

Two key issues in solvent design are solvent compatibility with downstream operations and plant-wide environmental considerations. Only on a global process basis can the selection of solvents lead to consistent, cost optimal, and environmentally benign results. In order to take account of these issues in a systematic way, Pistikopoulos and Stefanis (1998) extended their Methodology for Environmental Impact Minimization (MEIM) (Pistikopoulos et al., 1994) and applied it to optimal solvent design for single solvents.

Since solvents must satisfy cost, processing, environmental, safety and health related specifications, solvent selection is a very complex problem. MEIM provides the framework to tackle problems of this type, and is a design methodology for the assessment and minimization of the environmental impact of process systems. It relies on principles of life cycle analysis, which are embedded within a formal process optimization framework. In this way, environmental objectives are considered along with economics at the design stage, so as to determine cost-effective solutions to waste minimization problems.

In this work we show how the solvent design procedure within MEIM can be extended for the systematic identification and selection of blends of environmentally benign solvents for nonreactive, multicomponent absorption. The proposed developments include new property and operational feasibility constraints, a refined task based problem formulation, and a new solution procedure.

Motivating Example

Consider an industrial flowsheet for reduction of acrylonitrile (ACN) and dimethyl-formamide (DMF) emissions from an acrylic fiber manufacturing process (Pistikopoulos and Stefanis, 1998). As illustrated in Figure 1, gaseous effluents from the acrylic fiber plant are treated in two consecutive separation steps. Vapors rich in ACN and DMF coming from the main plant are mixed with those coming from the head of the packed column C-1701 and fed to a second packed column which uses very large quantities of demineralized water as a solvent.

The contaminated water is fed to an existing water treatment system. Although selective towards DMF, water exhibits poor selectivity for ACN, so that the ACN levels in the treated gas leaving column C-1702 are unacceptably high. However, the cost of piping this gas to the closest flare is prohibitive, so that the vapors are released directly to the

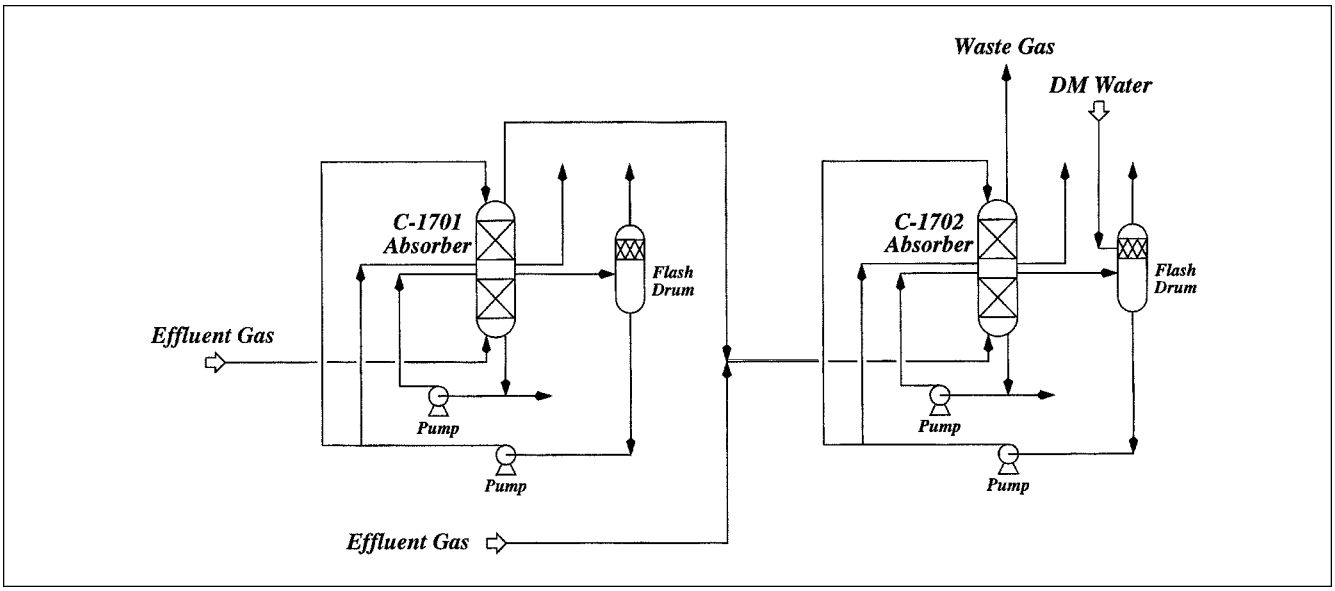

Figure 1. Gas treatment.

Table 1. Environmental Damage and Cost of Existing Gas Treatment System

ACN Emissions (ppm)	60
Process CTAM (ton air/h)	1,170
Treatment Cost (M$/yr)	4.1

environment. ACN is subject to a strict urban concentration limit and emissions incur a high environmental penalty cost (4 ppm and $60/ppb y, respectively). Thus, the existing process exhibits very poor economic and environmental performance, as presented in Table 1.

Several retrofitting attempts on the design and operation of the main process retaining water as solvent have been unsuccessful, due to the poor selectivity of water towards ACN. Clearly then, the process is a candidate for solvent substitution by which we should be able both to increase separation selectivity and reduce solvent flow. Replacement of the solvent water with an organic compound necessitates a solvent regeneration column with an associated capital cost. However, an alternative is to introduce into the water an additional solvent which is compatible with the existing water treatment and disposal system. This approach incurs only the cost of the continual consumption of this solvent.

This example gives rise to the following issues regarding the importance of solvent selection and its effect on the environment and plant economics:

- In multicomponent absorption, a single solvent component may not exhibit the appropriate properties which are necessary to satisfy the processing and environmental constraints; a solvent blend may be more appropriate and more economical.
- The choice of the solvent affects the environmental and economic performance of the separation task; a systematic procedure for solvent selection is required to improve the process operational and environmental performance on a task level.
- Solvent alternatives ought to be tested in terms of their process and global environmental and economic performances; a process wide verification is necessary to account for difficulties associated with solvent regeneration while global plant-wide verification is necessary to account for additional waste inputs to the process such as the pollution associated with meeting plant energy requirements.

We address these issues by extending the solvent design procedure of Pistikopoulos and Stefanis (1998) to allow the design of solvent blends for nonreactive, multicomponent absorption. This procedure represents an application of the Methodology for Environmental Impact Minimization (MEIM) (Pistikopoulos et al., 1994) which is summarized below.

MEIM: Brief Overview

The methodology for environmental impact minimization is a systematic tool for the estimation and minimization of the range of adverse effects of processing systems on the environment. This tool includes the following steps: the definition of the process system boundary, the environmental impact assessment of the emissions of the system and the synthesis of environmentally benign processes.

The methodology is flexible with respect to the choice of system boundary, which may be local or global so as to capture not only the conventional process emissions but also those associated with the manufacture of inputs (raw materials, capital, and energy) and the use of outputs.

The environmental impact assessment involves the compilation of an emissions inventory for the process of interest—the vector of gaseous, liquid and solid waste disposed to the environment from within the defined system boundary—and the transformation of this inventory into an impact vector. The emissions inventory, typically of high dimensionality, relies only on the mass of pollutant discharged and gives no indication of the form or extent of the actual damage that the emissions inflict on the environment. In order to reduce the high dimensionality of this vector and measure the environmental burden of the releases, the inventory is transformed into an impact vector by aggregating the emissions into impact metrics according to the form of the environmental burden caused (Pistikopoulos et al., 1994). Both routine and nonroutine releases may be treated in this way (Stefanis and Pistikopoulos, 1997).

As illustrated in Table 2, the initial waste vector is transformed into a condensed environmental impact vector that consists of metrics to measure pollution related to energy, global warming, air pollution, water pollution, and so on. In general, routine and nonroutine environmental damage can be classified as short or long term. Short-term environmental assessment is the measurement of damage at the point source of release (to air, water or soil). Long-term environmental assessment includes both measurement of the modification of the environment (such as through global warming and so on) and determination of pollutant fate, accounting for pollutant reactions and interactions in order to estimate the ultimate environmental impact of a set of emissions (Pistikopoulos and Stefanis, 1998). In this work we consider the short-term impact for a continuous, routine release scenario.

Table 2. Transformation of Emissions Inventory to Environmental Impact

Initial Vector	Condensed Vector
Energy contents of feedstocks and byproducts Processing energy Transport energy	Primary energy
C_1s, C_2s C_3s C_4s Other volatile organic compounds	Indirect global warming Photochemical oxidation
CO_2, CH_4 CFCs, N_2O HCFCs, CCl_4, CH_3CCl_3	Global warming (direct)
SO_2	Acid rain
NH_3, CO NO_x HCl, SO_2	Toxic air pollutants
Acids Heavy metals Dissolved, suspended solids BOD	Toxic water pollutants
Solid wastes	Solid wastes

The last step, which constitutes the heart of the methodology, is the incorporation of the environmental impact criteria into an overall process synthesis and optimization strategy. The process synthesis problem then conceptually involves determining the best design and plant operation featuring minimum environmental impact at minimum annualized cost on a plant wide basis. Formal multiobjective and/or parametric optimization techniques (see, for example, Acevedo and Pistikopoulos, 1996) can be applied to certain classes of this problem to obtain the Pareto space of (parametric) optimal solutions with respect to cost and the various components of environmental impact.

A large proportion of chemical plant emissions originate from the separation block, so that the selection of separation agents may have a profound effect on the environmental performance of a process. Although these agents could conceptually be considered as distinct design alternatives, in many cases there exists a very wide variety of candidates.

To address these issues, MEIM was extended to include principles of computer-aided molecular design (Pistikopoulos and Stefanis, 1998), in which a set of structural groups are systematically combined to form molecules with desired environmental and processing properties. In particular, the third step of the methodology was decomposed into a step-wise procedure which facilitates the selection of environmentally benign solvents while at the same time improving the separation efficiency and helping to meet plant-wide environmental and operational targets. In this article, we further extend this approach to the design of solvent blends.

Optimal Solvent Blend Design for Minimum Environmental Impact

In order to account for ecological considerations in solvent blend design, the following problem is addressed here:

Given

a processing manufacturing scheme with specified reaction routes, separation/purification alternatives and a set of constituent organic groups

the objective is

to identify the best blend of feasible agent molecules with optimal separation performance while simultaneously minimizing the adverse effects of the process on the environment at reduced cost.

The following three step procedure is applied (Pistikopoulos and Stefanis, 1998), involving:

- the identification of agent-based operations within the process of interest (such as gas absorption, liquid extraction) and specification of performance constraints for each separation task,
- at the separation task level, the determination of a list of candidate solvent *blends* satisfying processing and environmental constraints and,
- at the process level, the selection of an optimal solvent blend (from the list) based on global plant-wide process and environmental constraints.

Such a step-wise procedure, detailed in the subsequent sections, is particularly suitable since it reduces the combinatorial complexity and avoids unnecessary computations.

Solvent Blend Identification Formulation: Separation Task Level

Overview

The nonreactive, multicomponent absorption solvent blend identification problem can be defined as follows. Given, (i) a multicomponent stream of fixed flow rate and composition, (ii) an absorption task with known operating conditions, (iii) a set of structural groups (such as $-CH_3$, $-CHO$ ...) and their associated group contribution parameters, (iv) the type of agent molecule required (acyclic, monocyclic, or bicyclic), and (v) a minimum composition approach, for each pair of process/agent streams, so as to ensure feasible mass transfer, then the objective is to design a blend of feasible agent molecules featuring optimal separation performance and minimum environmental impact. In the analysis, we assume constant inlet/outlet stream mass flow rates, isobaric operation at a known pressure P_{tot}, 100% pure and regenerable solvents, no reaction occurring in the separation tasks, negligible enthalpies of solution and homogeneous mixtures. We also accept the inherent inaccuracies of the many property prediction techniques and the thermodynamic models we employ since, while the influence of parameter uncertainty on property prediction is an important issue, it is beyond the scope of the current work.

Our formulation is based on that reported by Pistikopoulos and Stefanis (1998), however, we include some additional property equations and operational feasibility constraints. Furthermore, we introduce a reformulation of the UNIFAC equations and a new solution procedure.

The model consists of five sets of equations: a set of structural and chemical feasibility constraints for the solvent molecules, pure component physical and environmental property prediction equations, nonideal multicomponent vapor-liquid equilibrium equations (reformulated UNIFAC equations), mixture property based operational feasibility constraints, and our process model. The sets employed in the model are shown in Table 3. All equations are detailed below, followed by a description of our solution technique.

Structural and chemical feasibility constraints

According to Odele and Macchietto (1993), if a collection of groups is to form a *structurally* feasible molecule, two constraints must be obeyed. The octet rule ensures that solvent s, made up of groups j, has zero valency (that is, there are no remaining free attachments between the groups)

$$\sum_j \left(2-\nu_j\right) n_{sj} = 2m, \quad \forall s \in S \tag{1}$$

where n_{sj} is the number of times group j occurs in solvent s, ν_j is the valency (the number of free attachments) of group j

Table 3. Task Based Model Sets

I	Set of all molecules i
ii	Alias for i
$S\ (\subset I)$	Solvent molecules
$NS\ (\subset I)$	Nonsolvent molecules
J	Set of all structural groups j
$nj,\ mj$	Aliases for j

and $m = 1$, 0 or -1 for acyclic, monocyclic, and bicyclic compounds, respectively. Equation 2 ensures that no two adjacent groups are linked by more than one bond (Odele and Macchietto, 1993) and, furthermore, that only one molecule is formed from any set of groups

$$\sum_j n_{sj} \geq n_{sj}(\nu_j - 1) + 2, \quad \forall s \in S,\ j \in J \tag{2}$$

Note that these rules do not imply unique molecular connectivity; for group sets of four or more, there may be several possible connectivities. In this case each connectivity should be considered as a separate candidate in the verification exercise.

An additional set of constraints is necessary to ensure *chemical* feasibility in the generated molecules and to limit the extent of the optimization search space. A set of such constraints were derived for the group sets selected in our example problems by studying Archer (1996) who provides a list of 289 industrial solvents. The constraints which were developed, according to the listed solvents, preclude certain combinations of groups, limit the occurrences of certain groups with certain others and provide upper and lower bounds on the number of functional and secondary groups which may appear within the generated solvent molecules.

Constraints developed in this way are empirical and are not based on chemical properties. However, they are both reasonable and appropriate for our example problems. Similar sets of rules for other applications may be readily developed by consulting lists and tables of molecular physical and thermodynamic properties which are widely available in the open literature. More general rules encompassing a wider range of groups are available, for example, in Constantinou et al. (1996), although the origin of these rules has never been published.

In order to write our chemical feasibility constraints, we introduce the binary (zero/one) variable INT_{sjt}, where t is the number of times group j occurs in solvent s. t runs from zero to t_{MAX} and so provides an additional upper limit on the occurrences of all groups in each solvent. INT_{sjt} takes the value unity if group j occurs t' times in solvent s, and zero for *all* other values of $t \neq t'$. INT_{sjt} is related to n_{sj} in the following way

$$n_{sj} = \sum_t t \cdot \mathrm{INT}_{sjt}, \quad \forall s \in S,\ j \in J \tag{3}$$

$$\sum_t \mathrm{INT}_{sjt} = 1, \quad \forall s \in S,\ j \in J \tag{4}$$

The chemical feasibility constraints are written in terms of both n_{sj} and INT_{sjt}, as shown below. These equations are provided for illustration only, and a full list of the chemical feasibility constraints derived from Archer is provided in Appendix A.

- The maximum and minimum numbers of groups are limited as follows

$$\sum_j n_{sj} \leq 9, \quad \forall s \in S$$

$$\sum_j n_{sj} \geq 1, \quad \forall s \in S$$

- The maximum number of functional groups is limited to four

$$n_{s,\mathrm{CH_3COO}} + n_{s,\mathrm{CH_2COO}} + n_{s,\mathrm{CH_3O}} + n_{s,\mathrm{CH_2O}} + n_{s,\mathrm{COCH_3}} + n_{s,\mathrm{COCH_2}} + n_{s,\mathrm{CHO}} + n_{s,\mathrm{OH}} \leq 4, \quad \forall s \in S$$

- The hydroxyl group may appear up to three times

$$n_{s,\mathrm{OH}} \leq 3, \quad \forall s \in S$$

- Groups $\mathrm{CH_2COO}$, $\mathrm{CH_3O}$ and $\mathrm{CH_2O}$ never appear together

$$\sum_t \left(\mathrm{INT}_{s,\mathrm{CH_2COO},t} + \mathrm{INT}_{s,\mathrm{CH_3O},t} + \mathrm{INT}_{s,\mathrm{CH_2O},t}\right) \leq 1, \quad \forall s \in S$$

The need for INT_{sjt} is clear if we consider this last equation. If we were to substitute n_{sj} for INT_{sjt}, we would not only preclude the occurrence of any of the three groups with the other two, we would also limit the occurrence of each group to at most once in any solvent. INT_{sjt} is also used to develop integer cuts as part of our solution procedure (see the subsection on solution procedure).

Note that we do not impose any constraints to prevent the same solvent appearing more than once in any solution. Thus, although the number of solvents to be selected is fixed by the dimension of the set S, some of the generated molecules may be the same. In this way, we do not force the generation of blends. The algorithm is flexible to select any number of *different* solvents subject to the maximum imposed by set S and if a single solvent solution satisfies all constraints, it is not excluded.

Note also that in addition to structural and chemical feasibility constraints, efficient group preselection rules (such as those proposed by Jaksland and Gani, 1996) can reduce the computational size of the problem and prevent selection of suboptimal molecules.

Pure component physical and environmental property equations

A key element of task-based solvent blend identification is the determination of the important pure component and mixture properties for the task and the formulation of appropriate property constraints. Many properties may be employed in the design of absorption solvents. In this section we discuss the pure component properties that we have selected for our model and list the property prediction methods we use. Mixture properties and the associated operational feasibility constraints are dealt with in the subsequent sections.

The solvent molecules are not known in advance, and we construct them from groups. Group contribution techniques provide direct access to many of the pure component properties we need using molecular structure information only. However, many pure component properties depend not only on structure but also on at least one other factor (such as temperature). Such properties may be predicted using correlations. Thus, we employ a mixture of group contribution techniques and equation oriented property prediction approaches.

Since the nonsolvent molecules are known in advance, we are sometimes able to estimate their properties through much simpler routes, or even enter them as parameters in some cases. This minimizes the computational complexity of property prediction.

Vapor Pressure. Pure component vapor pressure is central in determining the vapor-liquid partitioning of our components. For the solvent molecules, we use the Cox-Antoine relationship, from Reid et al. (1987), which gives reduced vapor pressure $P_{s,\mathrm{red}}^{\mathrm{vap}}$ as a function of the boiling point, critical temperature and pressure of solvent s with accuracy $< \pm 10\%$.

$$\log P_{s,\mathrm{red}}^{\mathrm{vap}} = \left(\frac{T_s^b - C_s}{T_s^{cr} - T_s^b}\right)\left(\frac{T - T_s^{cr}}{T - C_s}\right)\log P_s^{cr}, \quad \forall s \in S \quad (5)$$

where

$$C_s = -18 + 0.19 T_s^b, \quad \forall s \in S \quad (6)$$

For the solvents, low vapor pressure is desirable, since this implies low solvent loss. Therefore, we include an upper bound for the solvent vapor pressures in our model

$$P_s^{\mathrm{vap}} \le P_s^{\mathrm{vap,up}}, \quad \forall s \in S \quad (7)$$

For the nonsolvent molecules, we employ the appropriate form of the Antoine equation, again from Reid et al. (1987). The solvent normal boiling points and critical parameters are predicted using the following methods.

Normal Boiling Point. The group contribution method of Joback (1984) is employed to estimate solvent boiling points T_s^b (K)

$$T_s^b = 198 + \sum_j n_{sj}\delta_j^b, \quad \forall s \in S \quad (8)$$

where δ_j^b is the contribution of group j from tables. To facilitate separation in solvent regeneration and to avoid the formation of azeotropes, the boiling points of the solvents should be significantly higher, or lower, than those of the nonsolvent molecules that we wish to absorb. Furthermore, low solvent boiling points imply high solvent loss, whereas high solvent boiling points imply high energy demand in regeneration. Therefore, we bound the solvent boiling points as follows

$$T_s^{b,\mathrm{lo}} \le T_s^b \le T_s^{b,\mathrm{up}}, \quad \forall s \in S \quad (9)$$

where

$$T_s^{b,\mathrm{lo}} \ge T_{ns}^{b,\max} + \epsilon, \text{ or } T_s^{b,\mathrm{up}} \le T_{ns}^{b,\min} - \epsilon, \quad \forall s \in S \quad (10)$$

where ϵ is an adjustable minimum approach temperature parameter. The normal boiling points of the nonsolvent molecules are entered as parameters.

Critical Properties. We use the critical parameters in predicting not only solvent vapor pressures, but also the liquid molar volumes, heats of vaporization, and heat capacities for *all* materials. Nonsolvent critical parameters are entered as parameters. Based on Joback (1984) and Horvath (1992), the critical temperature T^{cr} (K), pressure P^{cr} (bar), and volume V^{cr} (cm^3/mol) for solvent s are estimated using the following group contribution expressions

$$\frac{T_s^b}{T_s^{cr}} = 0.584 + 0.965\sum_j n_{sj}\delta_j^T - \left(\sum_j n_{sj}\delta_j^T\right)^2, \quad \forall s \in S \quad (11)$$

$$P_s^{cr} = \frac{1}{\left(0.133 + 0.0032\, n_s^{\mathrm{atom}} - \sum_j n_{sj}\delta_j^P\right)^2}, \quad \forall s \in S \quad (12)$$

$$V_s^{cr} = 17.5 + \sum_j n_{sj}\delta_j^V, \quad \forall s \in S \quad (13)$$

where δ_j^T, δ_j^P, δ_j^V are the group contributions for group j and n_s^{atom} is the number of atoms in the molecule s.

Compressed Liquid Molar Volume. To maintain a homogeneous mixture and prevent phase separation, the densities of the solvents should not be widely different. We estimate the solvent densities through the compressed volumes.

According to Reid et al. (1987), Hankinson and Thomas (1979) present the following correlation for *saturated* liquid density V_i^{sat} (cm^3/mol)

$$\frac{V_i^{\mathrm{sat}}}{V_i^*} = V_{i,R}^0\left(1 - \omega_{\mathrm{SRK}} V_{i,R}^{\delta}\right), \quad \forall i \in I \quad (14)$$

$$V_{i,R}^0 = 1 + a\left(1 - T_i^{\mathrm{red}}\right)^{1/3} + b\left(1 - T_i^{\mathrm{red}}\right)^{2/3} + c\left(1 - T_i^{\mathrm{red}}\right) + d\left(1 - T_i^{\mathrm{red}}\right)^{4/3} \quad \forall i \in I,\ 0.25 < T_i^{\mathrm{red}} < 0.95 \quad (15)$$

$$V_{i,R}^{\delta} = \frac{\left(e + fT_i^{\mathrm{red}} + g\left(T_i^{\mathrm{red}}\right)^2 + h\left(T_i^{\mathrm{red}}\right)^3\right)}{\left(T_i^{\mathrm{red}} - 1.0001\right)}, \quad \forall i \in I \quad 0.25 < T_i^{\mathrm{red}} < 1 \quad (16)$$

The values of the constants for Eqs. 15 and 16 are given in Table 4. V_i^* is a component characteristic volume which may be estimated as follows

$$V_i^* = \frac{RT_i^{cr}}{P_i^{cr}}\left(a + b\omega_{\mathrm{SRK}} + c\omega_{\mathrm{SRK}}^2\right), \quad \forall i \in I \quad (17)$$

Values of the constants in Eq. 17 are given for several classifications of compounds in Reid et al. (1987).

Alternatively, V_i^* may be replaced by the critical volume with a resulting error in V_i^{sat} of up to 4%. ω_{SRK} is the acentric factor which forces the Soave equation to give a best fit of existing vapor pressure data. This may be replaced by the true acentric factor if necessary, again with an associated error of up to 4% in V_i^{sat}.

Thompson et al. (1982) extended this approach to the prediction of *compressed* liquid molar volumes (Reid et al., 1987)

Table 4. Constants for Saturated Liquid Density Equations

a	−1.52816	b	1.43907	c	−0.81446	d	0.190454
e	−0.296123	f	0.386914	g	−0.0427258	h	−0.0480645

Table 5. Constants for Compressed Liquid Density Equations

a	−9.070217	b	62.45326	d	−135.1102	f	4.79594
g	0.20047	h	1.14188	j	0.0861488	k	0.0344483

$$V_i = V_i^{\text{sat}}\left[1 - c\ln\left(\frac{\beta + P}{\beta + P_i^{\text{vap}}}\right)\right], \quad \forall i \in I \tag{18}$$

where

$$\frac{\beta}{P_i^{cr}} = -1 + a\left(1 - T_i^{cr}\right)^{1/3} + b\left(1 - T_i^{cr}\right)^{2/3} + d\left(1 - T_i^{cr}\right) + e\left(1 - T_i^{cr}\right)^{4/3}, \quad \forall i \in I \tag{19}$$

$$e = e^{(f + g\omega_{\text{SRK}} + h(\omega_{\text{SRK}})^2)}, \quad \forall i \in I \tag{20}$$

$$c = j + k\omega_{\text{SRK}}, \quad \forall i \in I \tag{21}$$

The constants for Eqs. 18–21 are given in Table 5. We bound the solvent *densities* as follows, where mw_s is the solvent molecular weight, in g/mol

$$\rho_s^{\text{lo}} < \frac{mw_s}{V_s} < \rho_s^{\text{up}}, \quad \forall s \in S \tag{22}$$

Enthalpy. In order to write energy balances within our process model, we must estimate the enthalpies of our process streams and the enthalpy changes associated with vaporization and condensation. We first calculate pure component liquid and vapor enthalpies relative to 273.15 K from

$$H_i^{\text{liq}}(T) = \int_{T_{\text{ref}}}^{T} C_{p,i}^{\text{liq}}\, dT, \quad \forall i \in I \tag{23}$$

$$H_i^{\text{vap}}(T) = \int_{T_{\text{ref}}}^{T_i^b} C_{p,i}^{\text{liq}}\, dT + \Delta H_i^{\text{vap}}\left(T_i^b\right) + \int_{T_i^b}^{T} C_{p,i}^{\text{gas}}\, dT, \quad \forall i \in I \tag{24}$$

The enthalpy of the mixture for each stream is then approximated by

$$H_{\text{liq}} = \sum_i H_i^{\text{liq}} X_i^l$$

$$H_{\text{vap}} = \sum_i H_i^{\text{vap}} X_i^l$$

Heat Capacity. We estimate the ideal gas molar heat capacities $C_{p,i}^{\text{gas}}$ (J/mol·K) and liquid molar heat capacities $C_{p,i}^{\text{liq}}$ (J/mol·K) of *all* materials using the following polynomial equations

$$C_{p,i}^{\text{gas}} = \left(\sum_j n_{ij}\Delta_a^j - 37.93\right) + \left(\sum_j n_{ij}\Delta_b^j + 0.21\right)T + \left(\sum_j n_{ij}\Delta_c^j - 3.91\times 10^{-4}\right)T^2 + \left(\sum_j n_{ij}\Delta_d^j + 2.06\times 10^{-7}\right)T^3, \quad \forall i \in I \tag{25}$$

$$\frac{C_{p,i}^{\text{liq}} - C_{p,i}^{\text{gas}}}{R} = 1.45 + 0.45\left(1 - T_i^{\text{red}}\right)^{-1} + 0.25\times\omega_s \times\left[17.11 + 25.2\left(1 - T_i^{\text{red}}\right)^{0.33} T_i^{\text{red}\,-1} + 1.742\left(1 - T_i^{\text{red}}\right)^{-1}\right], \quad \forall i \in I \tag{26}$$

where $T_{\text{red},i}$ is the reduced temperature of each compound i and the coefficients Δ_a^j, Δ_b^j, Δ_c^j and Δ_d^j are group contribution parameters from Joback and Stephanopoulos (1989). ω_i denotes the acentric factor of molecule i, which may be estimated from the vapor pressure of molecule i at a reduced temperature of 0.7 (Horvath, 1992)

$$\omega_i = -\log_{10} P_{\text{red},i}^{\text{vap}}\left(T_i^{\text{red}} = 0.7\right) - 1.0, \quad \forall i \in I \tag{27}$$

Ideal gas molar heat capacities for the known nonsolvent molecules are available directly as temperature-dependent polynomials (Reid et al., 1987). However, there is virtually no computational advantage in using these as opposed to the above group contribution methods.

Heat of Vaporization. The Chen equation (Horvath, 1992) provides a good estimate of solvent heats of vaporization $\Delta H_{s,298}^{\text{vap}}$. This relationship yields only 1.82% average deviation for many organic compounds. It employs the critical parameters and normal boiling point

$$\Delta H_{i,298}^{\text{vap}} = \frac{T_i^b\left[7.11\log P_i^{cr} - 7.82 + 7.9\dfrac{T_i^b}{T_i^{cr}}\right]}{1.07 - \dfrac{T_i^b}{T_i^{cr}}}, \quad \forall i \in I \tag{28}$$

Latent heat of vaporization decreases steadily with temperature and reaches zero at the critical point. The Watson relation (from Reid et al., 1987) provides a correlation between ΔH_i^{vap} and T_{oper}

$$\Delta H_i^{\text{vap}}(T_2) = \Delta H_i^{\text{vap}}(T_1)\left(\frac{1 - T_{i,1}^{\text{red}}}{1 - T_{i,2}^{\text{red}}}\right)^n, \quad \forall i \in I \tag{29}$$

where a common choice for n is 0.38. High heat of vaporization implies high energy demand in regeneration, so we impose an upper bound as follows

$$\Delta H_s^{\text{vap}}(T) \le \Delta H_s^{\text{vap,up}}(T), \quad \forall s \in S \tag{30}$$

Nonsolvent heats of vaporization at 298 K are entered as parameters, but are modified for temperature according to Eq. 29.

Lethal Concentration. In order to measure the short-term environmental impact of any materials released, we estimate the toxicity of each compound i using the group contribution technique of Gao et al. (1992). We use these toxicities as effective standard limit values (SLVs) for the emissions

$$-\log(LC50_i) = \sum_j n_{ij}\alpha_j, \quad \forall i \in I \tag{31}$$

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

Solubility Parameters. Solubility parameters provide a good indication of the compatibility of a solvent with a given solute. According to Hildebrand and Scott (1950), for the solution process to occur, the solubility parameter of the solvent must be nearly equal to that of the solute. Conversely, when there is a disparity between the two values, poor solubility is predicted.

According to Hansen (1967), the total solubility parameter δ_i^t of material i may be decomposed into three contributions: the dispersive solubility parameter δ_i^D, the polar solubility parameter δ_i^P, and the hydrogen bonding solubility parameter δ_i^H, as follows

$$\left(\delta_i^t\right)^2 = \left(\delta_i^D\right)^2 + \left(\delta_i^P\right)^2 + \left(\delta_i^H\right)^2$$

These individual parameters may be used to identify a more accurate solvent/solute match (Klein et al., 1992). They are related to the activity coefficient of material i in the following way (Lo et al., 1983)

$$\ln \gamma_i = \ln V_i - \ln \overline{V} + \left(1 - \frac{V_i}{\overline{V}}\right) + \frac{V_i}{RT}\left\{\left(\delta_i^D - \overline{\delta}^D\right)^2 + A_i\left[\left(\delta_i^P - \overline{\delta}^P\right)^2 + \left(\delta_i^H - \overline{\delta}^H\right)^2\right]\right\}$$

where V_i is the molar volume of component i; $\overline{V}$ is the mole fraction average molar volume; δ_i^D, δ_i^P, and δ_i^H are the dispersive, polar forces and hydrogen bonding solubility parameters of component i, respectively; A_i is a scaling factor and $\overline{\delta}^D$, $\overline{\delta}^P$ and $\overline{\delta}^H$ are the molar fraction averaged molar volume averages of the dispersive, polar forces and hydrogen bonding solubility parameters, respectively. For example

$$\overline{\delta}^D = \frac{\sum_i x_i V_i \delta_i^D}{\sum_i x_i V_i}$$

Here, we see quite clearly that similar solubility parameters contribute to lower activity coefficients, and, thus, indicate compatibility between solute and solvent.

Group contribution estimation techniques are available for each solubility parameter (Small, 1953). However, according to Joback (1989), they are neither linear nor linearizable. Thus, using solubility parameter data from Barton (1983) for 77 compounds, Joback developed his own *linear* group contribution estimation techniques. He argues that for most liquid systems, the differences between the dispersive solubility parameters are much smaller than the differences between the polar and hydrogen bonding solubility parameters, so that dispersive solubility parameters are frequently ignored. Thus, he developed estimation techniques for δ_i^P and δ_i^H only.

Since Joback did not employ a UNIFAC group set, we have developed the following new linear group contribution estimation techniques, based on UNIFAC groups, using data for 193 compounds taken from Archer (1996). Although the range of dispersive forces solubility parameter values shown in Archer is indeed much smaller than the ranges shown for both polar forces and hydrogen bonding solubility parameters ($12.3 \le \delta_i^D \le 20.6$, $0 \le \delta_i^P \le 26.2$ and $0 \le \delta_i^H \le 29.3$), it is not insignificant. Thus, we include all three solubility parameters

$$\delta_i^D = \sum_j n_{i,j}\Delta_{j,\delta^D} + 16.8316, \quad \forall i \in I \tag{32}$$

$$\delta_i^P = \sum_j n_{i,j}\Delta_{j,\delta^P} + 3.4000, \quad \forall i \in I \tag{33}$$

$$\delta_i^H = \sum_j n_{i,j}\Delta_{j,\delta^H} + 9.8571, \quad \forall i \in I \tag{34}$$

We have derived contributions for 33 UNIFAC groups. Full details of these models, the group contributions, and the estimation errors are given in Appendix B.

We estimate the solubility parameters of all of our materials in this way. However, since the nonsolvent molecules are known in advance, we pre-calculate their parameters and use them to formulate bounds on the solvent solubility parameters, to ensure solvent/solute compatibility

$$\delta_{s,\text{lo}}^D \le \delta_s^D \le \delta_{s,\text{up}}^D, \quad \forall s \in S \tag{35}$$

$$\delta_{s,\text{lo}}^P \le \delta_s^P \le \delta_{s,\text{up},}^P \quad \forall s \in S \tag{36}$$

$$\delta_{s,\text{lo}}^H \le \delta_s^H \le \delta_{s,\text{up}}^H, \quad \forall s \in S \tag{37}$$

Pure component property based mixture property constraints

Pure component properties may be used to develop mixture property constraints if (or, if we assume that) there are no mixing effects. Under these conditions, the properties of a solvent mixture may be estimated as linear combinations of the properties of the pure solvents. For example, the dispersive solubility parameter of a solvent blend may be estimated as follows

$$\delta_{\text{mix}}^D = \sum_s X_s^{l_{\text{in}}} \delta_s^D$$

If the mixture property is bounded, such relationships provide mixture property based composition restrictions for a solvent blend. For many properties, for example, bubble point, this linearization represents a significant computational simplification.

In this work, linear mixing rules were used for computational simplicity. The impact of this depends on the materials which make up the solvent blend. If the materials are similar, so that deviations from ideal behavior are small across the composition range, linear mixing rules may be considered sufficiently accurate. However, when large deviations do occur, the use of such rules may compromise the optimal solutions. In such cases the verification exercise may be employed to check the mixture properties and adjust the blend compositions. Alternatively, full mixing rules may be used in place of their less accurate linear counterparts in the task based formulation, increasing the computational complexity.

Multicomponent vapor-liquid equilibrium equations

Fugacity and activity coefficients account for nonideality in multicomponent vapor-liquid equilibrium. The UNIWAALS method as reported in Gani et al. (1989) may be used as an equation of state to represent vapor (liquid) nonidealities and, hence, predict the fugacity coefficient ϕ_i for each component i as a function of composition, critical pressure, and temperature. However, in the examples presented, the fugacity coefficients are all approximated by unity due to the relatively low operating pressures (that is, we assume ideal gas behavior).

We employ the UNIFAC equations to predict activity coefficients. However, we do not use them in their original form; we reformulate them, leading to some significant simplifications. The original UNIFAC equations and the procedure of reformulation are detailed in the notation section and Appendix. Here we present the reformulated equations in their programming form. We have defined some new variables ($C1_i$, $R1_i$ and so on) which simply represent groupings of the original UNIFAC terms.

Overall Activity Coefficient Equation. The activity coefficient γ_i of component i, is made up of a combinatorial part γ_i^c and a residual part γ_i^r

$$\ln \gamma_i = \ln \gamma_i^c + \ln \gamma_i^r, \quad \forall i \in I \tag{38}$$

Combinatorial Part Equations. The combinatorial term takes into account entropy effects and depends on the size and shape of the molecule. It is essentially a complicated function of Q_i and R_i—the Van der Waals area and volume of component i, respectively

$$\ln \gamma_i^c = \ln R_i - \ln\left(\sum_{ii} X_{ii}^l R_{ii}\right) - 5Q_i \ln R_i - 5Q_i \ln\left(\sum_{ii} X_{ii}^l Q_{ii}\right) + 5Q_i \ln Q_i + 5Q_i \ln\left(\sum_{ii} X_{ii}^l R_{ii}\right) - 5Q_i + 1 + C1_i, \quad \forall i \in I \tag{39}$$

where

$$C1_i = \left(5\sum_{ii} X_{ii}^l Q_{ii} - 1\right)\left(\frac{R_i}{\sum_{ii} X_{ii}^l R_{ii}}\right), \quad \forall i \in I \tag{40}$$

Q_i and R_i are calculated simply by summing the appropriate Van der Waals group area and volume contributions q_j and r_j, respectively, obtained from Reid et al. (1987)

$$R_i = \sum_j n_{i,j} r_j, \quad \forall i \in I \tag{41}$$

$$Q_i = \sum_j n_{i,j} q_j, \quad \forall i \in I \tag{42}$$

Residual Part Equations. The residual term is mainly governed by the energetic interactions between different groups of molecules. It is itself a combination of two parts: a mixture term and a pure component term

$$\ln \gamma_i^r = R1_i - R2_i, \quad \forall i \in I \tag{43}$$

The mixture term $R1_i$ represents the summation of the group residual activity coefficients of all groups j from molecule i in the mixture with all other molecules present

$$R1_i = Q_i - \sum_j n_{i,j} q_j \ln R3_j + Q_i \ln\left(\sum_{ii} X_{ii}^l Q_{ii}\right) - \sum_j \left(\frac{R4_{i,j}}{R3_j}\right), \quad \forall i \in I \tag{44}$$

where

$$R3_j = \sum_{mj} q_{mj} \sum_{ii} n_{ii,mj} X_{ii}^l \psi_{mj,j}^{VLE}, \quad \forall j \in J \tag{45}$$

$$R4_{i,j} = \sum_{mj} n_{i,mj} q_{mj} q_j \sum_{ii} n_{ii,j} X_{ii}^l \psi_{mj,j}^{VLE}, \quad \forall i \in I, \quad \forall j \in J \tag{46}$$

The pure component term $R2_i$ represents the summation of the group residual activity coefficients of all groups j from molecule i as if in a reference solution containing only molecules of type i. This second term is necessary so that γ_i becomes unity as $X_i^l \to 1$

$$R2_i = Q_i \ln Q_i - \sum_j n_{i,j} q_j \ln\left(\sum_{mj} q_{mj} n_{i,mj} \psi_{mj,j}^{VLE}\right), \quad \forall i \in I \tag{47}$$

The interaction parameter for vapor/liquid equilibrium calculations $\psi_{nj,mj}^{VLE}$ is given by

$$\psi_{nj,mj}^{VLE} = e^{-(a_{nj,mj}^{VLE}/T)}, \quad \forall nj, mj \in J \tag{48}$$

$a_{nj,mj}^{VLE}$ represents the interaction parameter for group nj with group mj for vapor/liquid equilibrium calculations, reported in Reid et al. (1987). For dilute aqueous solutions of organic compounds, the interaction parameters reported by Chen et al. (1993) may be substituted for those provided by Reid. Any missing parameters are assigned the value of zero, as in Reid. Note that $a_{nj,mj}^{VLE} \neq a_{mj,nj}^{VLE}$, so that $\psi_{nj,mj}^{VLE} \neq \psi_{mj,nj}^{VLE}$.

To complete the UNIFAC reformulation, we must introduce the following condition

$$\sum_i X_i^l = 1 \tag{49}$$

This is necessary since we have eliminated several conditions from the original UNIFAC model (see Appendices C and D).

Mixture property operational feasibility constraints

Two important issues for multicomponent absorption systems are azeotropy and liquid phase stability (that is, miscibility). For a feasible solvent regeneration operation, we must avoid the existence of azeotropes, not only between the solvents and the nonsolvents, but also between the solvents themselves. Clearly, we also require complete mutual solubility for our solvent blend and for the mixture formed when

the nonsolvent molecules are introduced. Phase splitting may be detrimental to both absorption and regeneration operations and causes additional problems, such as pumping difficulties.

Determination of Binary Azeotropic Points. For a binary system, consisting of components *i* and *ii*, the following equilibrium relationships apply

$$Y_i = \frac{\gamma_i P_i^{\text{vap}} X_i}{\phi_i P^{\text{oper}}}$$

$$Y_{ii} = \frac{\gamma_{ii} P_{ii}^{\text{vap}} X_{ii}}{\phi_{ii} P^{\text{oper}}}$$

If an azeotrope exists, the following conditions apply at the azeotrope

$$Y_i^{az} = X_i^{az}$$

$$Y_{ii}^{az} = X_{ii}^{az}$$

so that

$$\frac{Y_i^{az}}{X_i^{az}} = \frac{Y_{ii}^{az}}{X_{ii}^{az}} = 1$$

Thus, substituting from the equilibrium expressions above and assuming ideal gas behavior so that the fugacity coefficients may be ignored, at the azeotrope we have

$$\gamma_i^{az} P_i^{\text{vap}} = \gamma_{ii}^{az} P_{ii}^{\text{vap}}$$

where the products γP^{vap} represent the effective vapor pressures exerted by the components in the mixture. This condition simply states that if an azeotrope exists, then at some point on the composition scale, the effective vapor pressures of the two components will be equal. For our purposes, it is not necessary to identify the exact composition of any azeotrope. It is sufficient to determine whether an azeotrope exists, so that we can avoid such mixtures. We achieve this by considering the effective vapor pressures at the extremes of the composition scale.

When the mixture comprises of component *i* with component *ii* at infinite dilution, the effective vapor pressures of components *i* and *ii* are P_i^{vap} and $\gamma_{ii,i}^{\infty} P_{ii}^{\text{vap}}$ respectively, where $\gamma_{ii,i}^{\infty}$ represents the activity coefficient of component *ii* at infinite dilution in component *i*. Conversely, when the mixture consists of component *ii*, with component *i* at infinite dilution, the effective vapor pressures are given by $\gamma_{i,ii}^{\infty} P_i^{\text{vap}}$ and P_{ii}^{vap}, respectively. If

$$P_i^{\text{vap}} > \gamma_{ii,i}^{\infty} P_{ii}^{\text{vap}} \quad \text{and} \quad \gamma_{i,ii}^{\infty} P_i^{\text{vap}} < P_{ii}^{\text{vap}},$$

or

$$P_i^{\text{vap}} < \gamma_{ii,i}^{\infty} P_{ii}^{\text{vap}} \quad \text{and} \quad \gamma_{i,ii}^{\infty} P_i^{\text{vap}} > P_{ii}^{\text{vap}},$$

then the effective vapor pressures must cross at some point on the composition scale. At this point, we have an azeotrope.

In binary mixtures, the effective vapor pressures of the two components vary monotonically with composition. Therefore, we can say that the first condition above indicates a minimum boiling point azeotrope, and the second a maximum boiling point azeotrope.

The two conditions may be combined into a single constraint to preclude the existence of binary azeotropes between our solvents and nonsolvents, and between the solvents themselves

$$\left(P_i^{\text{vap}} - \gamma_{ii,i}^{\infty} P_{ii}^{\text{vap}}\right)\left(P_{ii}^{\text{vap}} - \gamma_{i,ii}^{\infty} P_i^{\text{vap}}\right) < 0 \quad \forall i \neq ii,\ ii \in I \tag{50}$$

This test is used as a preliminary screening measure. Clearly, we cannot guarantee to avoid multicomponent azeotropes in this way. However, multicomponent azeotropes may be captured at the verification level using global optimization techniques (Harding et al., 1997).

In order to perform the azeotropy test, we estimate activity coefficients at infinite dilution using the following method.

Activity Coefficients at Infinite Dilution. We employ the UNIFAC equations to estimate activity coefficients at infinite dilution. For liquid/liquid equilibrium calculations, the tabulated UNIFAC interaction parameters are different from those used for vapor/liquid equilibrium calculations (Godfrey and Slater, 1994). LLE parameters, fitted by Magnussen (1981) are used in the equations below. Where LLE parameters are missing from Magnussen, VLE parameters (from Reid et al., 1987) are used instead, as suggested by Magnussen. Parameters which are missing altogether are assigned the value of zero, as in Reid.

For material *i* at infinite dilution in material *ii*, we can make the simplification

$$X_{ii}^{l} = 1$$

Thus, the UNIFAC equations reduce to the following.

Overall Activity Coefficient Equation.

$$\ln \gamma_{i,ii}^{\infty} = \ln \gamma_{i,ii}^{c,\infty} + \ln \gamma_{i,ii}^{r,\infty}, \quad \forall i \neq ii,\ ii \in I \tag{51}$$

Combinatorial Part Equations.

$$\ln \gamma_{i,ii}^{c,\infty} = \ln R_i - \ln R_{ii} - 5Q_i \ln R_i - 5Q_i \ln Q_{ii} + 5Q_i \ln Q_i + 5Q_i \ln R_{ii} - 5Q_i + 1 + C1_{i,ii}, \quad \forall i \neq ii,\ ii \in I \tag{52}$$

where

$$C1_{i,ii} = (5Q_{ii} - 1)\left(\frac{R_i}{R_{ii}}\right), \quad \forall i \neq ii,\ ii \in I \tag{53}$$

and as before

$$R_i = \sum_j n_{i,j} r_j, \quad \forall i \in I \tag{54}$$

$$Q_i = \sum_j n_{i,j} q_j, \quad \forall i \in I \tag{55}$$

Residual Part Equations.

$$\ln \gamma_{i,ii}^{r,\infty} = R1_{i,ii} - R2_i, \quad \forall i \neq ii, \; ii \in I \tag{56}$$

where

$$R1_{i,ii} = Q_i - \sum_j n_{i,j} q_j \ln R3_j + Q_i \ln Q_{ii} - \sum_j \left(\frac{R4_{i,ii,j}}{R3_{ii,j}} \right), \quad \forall i \neq ii, \; ii \in I \tag{57}$$

$$R3_{ii,j} = \sum_{mj} q_{mj} n_{ii,mj} \psi_{mj,j}^{LLE}, \quad \forall ii \in I, \; j \in J \tag{58}$$

$$R4_{i,ii,j} = \sum_{mj} n_{i,mj} q_{mj} q_j n_{ii,j} \psi_{mj,j}^{LLE}, \quad \forall i \neq ii, \; ii \in I, \; \forall j \in J \tag{59}$$

and

$$R2_i = Q_i \ln Q_i - \sum_j n_{i,j} q_j \ln \left(\sum_{mj} q_{mj} n_{i,mj} \psi_{mj,j}^{LLE} \right), \quad \forall i \in I \tag{60}$$

The interaction parameter $\psi_{nj,mj}^{LLE}$ is given by

$$\psi_{nj,mj}^{LLE} = e^{-(a_{nj,mj}/T)}, \quad \forall nj, mj \in J \tag{61}$$

Note that $a_{nj,mj}^{LLE} \neq a_{mj,nj}^{LLE}$, so that $\psi_{nj,mj}^{LLE} \neq \psi_{mj,nj}^{LLE}$.

Determination of Partial Miscibility. If a binary system, consisting of components *i* and *ii*, is to exhibit complete miscibility (phase stability), the following inequality must be satisfied across the entire composition range

$$\left[\frac{d^2 \Delta G}{d\left(X_i^l\right)^2} \right]_{T,P} > 0$$

where ΔG is the molar Gibbs energy change of mixing for the system. Smith et al. (1996) show that this inequality leads to the entirely equivalent phase stability condition

$$\left[\frac{d \ln \gamma_i}{dX_i^l} \right]_{T,P} > -\frac{1}{X_i^l}$$

which may be written in terms of the composition of either component. Although this condition applies in binary systems only, it may be introduced into the solvent design procedure in a similar way as the azeotropy condition to exclude solvents which form unstable mixtures either with the nonsolvent components or among themselves. However, since the condition must be satisfied everywhere, it cannot be reduced in the same way as the azeotropy condition or applied as a local test of phase stability. Thus, it should be applied across the composition range for each pair of materials as part of the verification exercise.

Phase stability for the entire *multicomponent* mixture may be investigated using global optimization approaches, such as that of McDonald and Floudas (1995), which determines the number of liquid phases through the minimization of the Gibbs Free Energy function.

Process model equations

At the task level, we consider only the process unit in which the solvent performs its separation task, in this case, the absorber. Our process model is based on the modular representation of Papalexandri and Pistikopoulos (1996), extended by Rahim Ismail et al. (1997), which provides an aggregated description of the separation task. The model also includes molecular performance constraints and environmental impact metrics, which we use to formulate our design objectives.

Mass/Heat Transfer Modular Representation. According to Papalexandri and Pistikopoulos (1996), a mass-/heat-transfer module as depicted in Figure 2, comprising of a combined mass-/heat-transfer unit and a pure heat-transfer unit, may be used to represent a separation task in an aggregated fashion. For an absorption process, only the combined mass-/heat-transfer unit with no bypass streams is required. This unit is described by a set of equations which govern the exchange of mass and/or heat between the adjacent streams, as follows.

Driving Force Constraints. Driving force constraints between adjacent streams are written for each component.

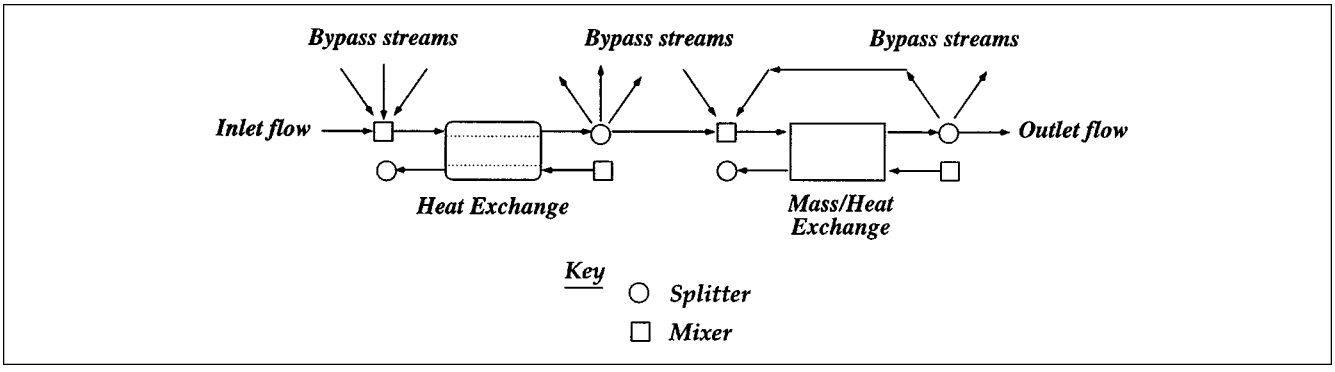

Figure 2. Mass-/heat-transfer module.

These describe the phase equilibrium limitation for mass transfer between solute rich and lean phases and are formulated as follows (Rahim Ismail et al., 1997).

When the composition of component i is above equilibrium *in the liquid*, we write the following driving force constraints

$$\mathcal{G}_1 \geq 0$$

$$\mathcal{G}_2 \geq 0$$

where

$$\mathcal{G}_1 = K_i^{l_{in}} X_i^{l_{in}} - X_i^{v_{out}} \tag{62}$$

$$\mathcal{G}_2 = K_i^{l_{out}} X_i^{l_{out}} - X_i^{v_{in}} \tag{63}$$

For the reverse case, when the composition of component i is above equilibrium *in the vapor*

$$\mathcal{G}_1 \leq 0$$

$$\mathcal{G}_2 \leq 0$$

In the general case, when the mass-transfer direction is to be determined, we have

$$\mathcal{G}_1 \mathcal{G}_3 \geq 0, \quad \forall i \in I \tag{64}$$

$$\mathcal{G}_2 \mathcal{G}_3 \geq 0, \quad \forall i \in I \tag{65}$$

where

$$\mathcal{G}_3 = F^{l_{in}} X_i^{l_{in}} - F^{l_{out}} X_i^{l_{out}}, \quad \forall i \in I \tag{66}$$

$K_i^{l_{in}}$ and $K_i^{l_{out}}$ are defined as follows in order to capture nonidealities in the vapor liquid equilibrium

$$K_i^{l_{in}} = \frac{\gamma_i^{l_{in}} P_i^{vap}}{\phi_i^{v_{out}} P_{tot}}, \quad \forall i \in I \tag{67}$$

$$K_i^{l_{out}} = \frac{\gamma_i^{l_{out}} P_i^{vap}}{\phi_i^{v_{in}} P_{tot}}, \quad \forall i \in I \tag{68}$$

Phase Defining Constraints. To define the liquid streams, an equation of the following form is introduced (Rahim Ismail et al., 1997)

$$\sum_i K_i X_i \leq 1 \tag{69}$$

where

$$K_i = \frac{\gamma_i P_i^{vap}}{\phi_i P_{tot}} \tag{70}$$

and for vapor streams

$$\sum_i \frac{X_i}{K_i} \leq 1 \tag{71}$$

Component Mass Balances. Around the mass/heat transfer block for each component i

$$F^{v_{in}} X_i^{v_{in}} + F^{l_{in}} X_i^{l_{in}} - F^{v_{out}} X_i^{v_{out}} - F^{l_{out}} X_i^{l_{out}} = 0, \quad \forall i \in I \tag{72}$$

where $F^{v_{in}}$, $F^{v_{out}}$ are the molar flow rates of the vapor stream at the inlet and outlet of the module respectively, $F^{l_{in}}$, $F^{l_{out}}$ are the molar flow rates of the liquid stream at the module inlet and outlet, and, similarly, $X^{v_{in}}$, $X^{l_{in}}$, $X^{v_{out}}$, and $X^{l_{out}}$ denote the inlet and outlet compositions of vapor and liquid streams. Additionally, for all streams around the mass/heat exchange block the summation of molar fractions must be equal to one. For example at the vapor inlet

$$\sum_i X_i^{v_{in}} = 1$$

Note we have already written this equation for the liquid as part of the reformulated UNIFAC model (see Eq. 49).

Energy Balance. The energy balance around the module takes the following form

$$F_i^{v_{in}} \overline{H}^{v_{in}} + F_i^{l_{in}} \overline{H}^{l_{in}} - F_i^{v_{out}} \overline{H}^{v_{out}} - F_i^{l_{out}} \overline{H}^{l-out} = 0 \tag{73}$$

where, for example, $\overline{H}^{v_{in}}$ is the average enthalpy of the vapor inlet stream calculated by

$$\overline{H}^{v_{in}} = \sum_i H_i^{v_{in}} X_i^{v_{in}} \tag{74}$$

where $H_i^{v_{in}}$ is the enthalpy of component i at temperature $T^{v_{in}}$.

Environmental Impact Metrics. Detailed environmental impact assessment criteria cannot be readily incorporated with molecular design principles at this stage. However, the availability of group contribution techniques for the prediction of environmental properties such as lethal concentration, LC50, and Ozone Depletion Potential allows the development of a reduced environmental impact vector, expressed in terms of molecular structure information, of the form

$$EI = [\text{CTAM CTWM SMD SODI}]_w^T$$

Metrics such as CTAM and CTWM can be quantified as a function of lethal concentration LC50 (mol/L) by ingestion or inhalation

$$\text{CTAM(CTWM)} = \sum_i \frac{F_i^o}{LC50_i} \tag{75}$$

Note that the idea of aggregating environmental damage provides a sensible way of determining environmentally benign molecules while simultaneously minimizing solvent and material losses. Hence, there is no need to formulate additional objectives such as solvent capacity and loss.

Molecular Performance Constraints. Based on the mass-/heat-transfer modular representation, targets are imposed on the loss of solvent materials and the capacity of the solvent blend to absorb the solute and so ensure feasible operating performance of the task

$$X_s^{v_{out}} \leq \text{Target Loss} \tag{76}$$

$$\frac{X_{ns}^{l_{out}}}{X_{ns}^{v_{out}}} \geq \text{Target Capacity} \tag{77}$$

We now wish to solve this entire model (Eqs. 1–77) in a multiobjective mixed integer optimization setting with environmental impact metrics (such as critical air mass, CTAM and critical water mass, CTWM) as the set of objective functions. The optimization variables are x, the set of continuous process variables (such as flow rates, temperatures, and mass fractions), and n_{ij}, the set of integer variables indicating the number of groups that constitute each molecule in the blend. Selected blends must meet performance standards imposed by constraints on material losses to the environment and solute absorption capabilities. In Pistikopoulos and Stefanis (1998), a similar problem was formulated for the design of single solvents but no solution procedure was put forward. To cope with the added complexity of solvent blend design, we have developed a new solution procedure which is described in the sequel.

Solution procedure

Overview. In order to solve the task-based solvent blend selection model we adopt a step-wise decomposition based algorithm which follows the general principles of the generalized benders decomposition (Geoffrion, 1974). We decompose the NLP (nonlinear programming) *primal* problem into seven steps which we solve in series in order to simplify and initialize the final primal NLP, as well as possible. Our solution procedure is shown in Figure 3.

In addition to the assumptions listed at the beginning of the section on solvent blend identification, we have assumed isothermal operation in the absorber process model. This allows us to extract a large number of temperature-dependent property prediction equations from the final NLP.

The energy balance is used to calculate the value that the temperature at the top of the absorber would take, assuming *only* that the bottom temperature is fixed at the existing op-

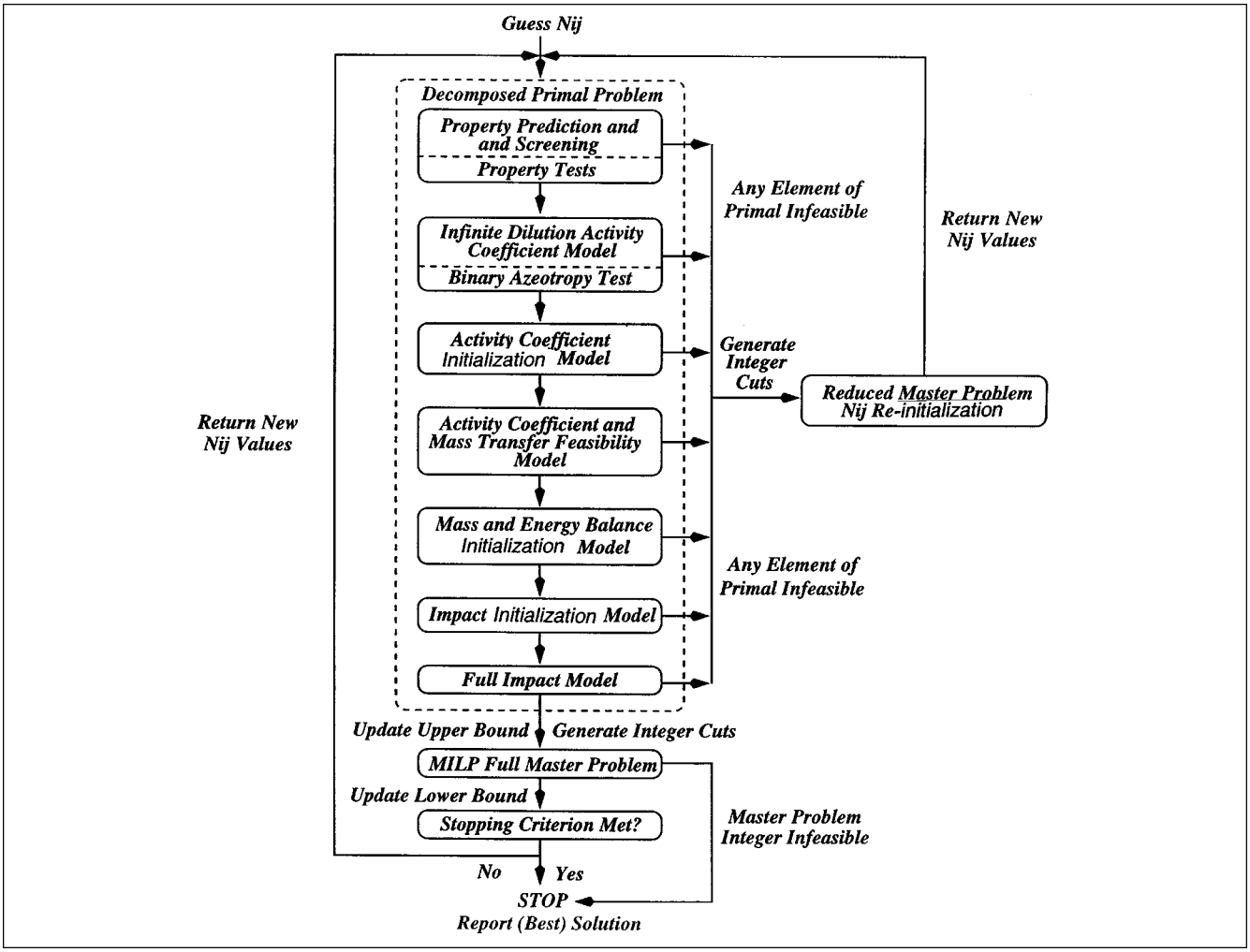

Figure 3. Task-based blend selection solution procedure.

erating temperature. The temperature change throughout the column allows us to judge the validity of the isothermal assumption. Clearly, we can explore solvent performance under different conditions by changing the operating temperature. The energy balance is solved within the decomposed primal NLP model.

Integer Initialization. The first step in the solution procedure is to provide an initial guess for the integer array n_{ij}, that is, to guess the identities of the solvents. Throughout the primal, the values n_{ij} do not change, so this guess is stored as a parameter array to be called as required.

Property Prediction Models and Tests. The first step of the decomposed primal problem involves property prediction and preliminary screening of the solvent molecules. We evaluate all the pure component properties according to the prevailing process conditions and test them against their bounds using all the Eqs. 5–37.

Since the operating temperature and pressure of the absorber are fixed, the values of all the pure component physical and environmental properties are also fixed throughout the primal NLP. Thus, these values may be stored as parameters to be called as required for the subsequent calculations, and we may remove Eqs. 5–37 from our final NLP, simplifying it significantly.

Furthermore, for systems involving only two solvents, we may use the pure component properties at this stage to determine the range of compositions which satisfy the pure component property-based mixture property constraints (discussed earlier). This is achieved using a linear optimization procedure which develops composition bounds for the solvent blend to be carried forward into the subsequent steps of the primal NLP. For systems involving more than two solvents, this is not possible, so the pure component property-based mixture property constraints are then included in the mass and energy balance initialization model and all subsequent NLP steps.

The second step of the primal involves the evaluation of the activity coefficients at infinite dilution using Eqs. 51–61 and the azeotropy test which is performed using Eq. 50. The infinite dilution activity coefficients are only needed to perform this test, so that these equations do not feature in the final NLP.

If there is property constraint violation, the azeotropy test is failed, or there is infeasibility in either of these NLP steps, the algorithm moves directly to a *reduced* mixed-integer linear programming (MILP) *master* problem which reinitializes the set of integers (that is, solvent molecules), and integer cuts are generated to exclude the offending combination of solvents (see below).

Activity Coefficient Models. The third and fourth primal steps are intended to find a feasible initial point for our reformulated UNIFAC equations (see the subsection on multicomponent vapor-liquid equilibrium equations), which is consistent with the process model. We begin by solving *only* the activity coefficient equations (Eqs. 38–49) with a dummy objective as follows

$$\text{minimize } u \tag{78}$$

$$OBJ_{\text{activity}} \le u \tag{79}$$

$$OBJ_{\text{activity}} \ge 0 \tag{80}$$

This step is designed to provide a feasible initial point for the reformulated UNIFAC equations. The next step is to integrate these equations with the process model. We note that the mass-transfer driving force constraints (Eqs. 64–66) provide the link between the UNIFAC and process models through the vapor/liquid equilibrium constants $K_i^{l_{\text{in}}}$ and $K_i^{l_{\text{out}}}$ (which depend on the activity coefficients predicted by UNIFAC—see Eqs. 67 and 68). Recognizing this, we formulate our model including *only* the mass-transfer driving force constraints and the reformulated UNIFAC equations. We exploit the fact that the driving force constraints are inequalities and introduce the slack variables $INF1_i$ and $INF2_i$

$$\mathcal{G}_1 \mathcal{G}_3 + INF1_i \ge 0, \quad \forall i \in I \tag{81}$$

$$\mathcal{G}_2 \mathcal{G}_3 + INF2_i \ge 0, \quad \forall i \in I \tag{82}$$

where, as before

$$\mathcal{G}_3 = F^{l_{\text{in}}} X_i^{l_{\text{in}}} - F^{l_{\text{out}}} X_i^{l_{\text{out}}}, \quad \forall i \in I$$

The slacks represent infeasibilities in the driving force equations, our objective is to minimize the sum of these infeasibilities

$$\text{minimize } \sum_i (INF1_i + INF2_i) \tag{83}$$

In this way, we develop a feasible initial point for the UNIFAC equations which is consistent with the process equations, using a reduced process model. Again, if there is infeasibility in either of these NLP steps, the algorithm moves directly to the reduced master, and integer cuts are automatically written to exclude the offending combination of solvents (see below).

Mass and Energy Balance Initialization Model. The fifth step of the primal involves only the mass and energy balance equations from the process model (Eqs. 64–74). We solve these equations using a dummy objective as before and the initial point developed above. As previously discussed, the energy balance is employed here to judge the validity of the isothermal assumption. For systems of more than two solvents, the pure component-based mixture property constraints are introduced here to bound the compositions of the solvent blend. As before, if there is infeasibility at this stage, the algorithm moves directly to the reduced master.

Impact Initialization and Full Impact Model. The impact initialization model draws together all the elements of the full impact NLP primal problem (the activity coefficient equations, the mass and energy balances, the impact metric equations, and the pure component-based mixture property constraints for systems of more than two solvents) except the objective. For initialization purposes, the model is solved using the initial points from above with a dummy objective.

The following objective is then introduced and the full impact NLP primal problem is solved for minimum impact

$$\text{minimize CTAM(CTWM)} \tag{84}$$

Values from the initialization model serve for initialization, however, here, all the continuous variables (except the

pre-determined property values) are free to take any values consistent with their bounds and the constraints. If there is infeasibility at this stage, the algorithm moves directly to the reduced master.

MILP Master Problems. We have two MILP (mixed-integer linear programming) *master* problems: a conventional one and a *reduced* master which reinitializes the set of solvent molecules in case of property test failure or infeasibility in any element of the decomposed primal problem.

The conventional master problem is made up of the pure integer structural and chemical feasibility constraints (Eqs. 1 and 2, and A1–A26) and an objective function which is constructed from all constraints in the full impact model which involve the array n_{ij}. This objective is written as follows (see Appendix E)

$$\text{minimize } \mu \tag{85}$$

where

CTAM(CTWM)

$$+\sum_i \lambda_i^a\left(R_i-\sum_j n_{i,j}r_j\right)+\sum_i \lambda_i^b\left(Q_i-\sum_j n_{i,j}q_j\right)$$
$$+\sum_i \lambda_i^c\left[R1_i-Q_i+\sum_j n_{i,j}q_j\ln R3_j\right.$$
$$\left.-Q_i\ln\left(\sum_{ii} X_{ii}^l Q_{ii}\right)+\sum_j\left(\frac{R4_{i,j}}{R3_j}\right)\right]$$
$$+\sum_i \lambda_i^d\left(R3_j-\sum_{mj} q_{mj}\sum_{ii} n_{ii,mj}X_{ii}^l\psi_{mj,j}^{VLE}\right)$$
$$+\sum_i \lambda_j^e\left(S1_j-\sum_{ii} n_{ii,j}X_{ii}^l\right)$$
$$+\sum_i\sum_j \lambda_{i,j}^f\left(S2_{i,j}-\sum_{mj} q_{mj}n_{i,mj}\psi_{mj,j}^{VLE}\right)$$
$$+\sum_i\sum_j \lambda_{i,j}^g\left(R4_{i,j}-\sum_{mj} n_{i,mj}q_{mj}q_jS1_j\psi_{mj,j}^{VLE}\right)$$
$$+\sum_i \lambda_i^h\left(R2_i-Q_i\ln Q_i-\sum_j n_{i,j}q_j\ln S2_{i,j}\right)\le\mu \tag{86}$$

$S1_j$ and $S2_{i,j}$ are additional variables defined in order to linearize the master objective (see Appendix E), and each λ represents a Lagrange multiplier taken from the solution of the full impact model. All continuous variables in Eq. 86 are fixed at the levels determined in the solution of the full impact model.

The reduced master is necessary since we have replaced the *relaxed* primal problem usually present in the GBD algorithm with a set of property tests and initialization models. If there is property test failure or any element of the primal is infeasible, the set of Lagrange multipliers necessary to run the full master will be incomplete, so we run a reduced master in its place. The reduced master involves *only* the structural and chemical feasibility constraints and a simplified objective which does not involve Lagrange multipliers. Recognizing that simple solvent molecules are more attractive than complex ones, the objective is designed to select the simplest remaining pair of solvents which satisfy all pure integer constraints

$$\text{minimize } \sum_s\sum_j n_{sj} \tag{87}$$

Before either master problem is run, integer cuts must be formulated to prevent previous solutions from recurring. The integer cuts are written in terms of the binary array INT_{sjt} to exclude previous *combinations* of solvents and not the individual solvents themselves. In this way, we allow rejected solvent molecules to reappear in different blends. The INT_{sjt} values needed to formulate the cuts are determined using Eqs. 3 and 4 as a pre-processing step before either master.

Convergence. We employ a conventional stopping (convergence) criterion. The upper bound is only updated if the full impact model reaches an optimal solution, and the lower bound is only updated if the *full* master reaches an optimal solution. Convergence is checked after every such iteration. If convergence has not been achieved (or the reduced master was employed), the series of NLPs is re-solved, using the new integer values from the master and the procedure is repeated until a solution is found. Should the master become integer infeasible, the algorithm stops and reports the current best solution.

In practice, this algorithm exhibits acceptable convergence behavior. Systematic initialization is achieved by the successive steps of the decomposed primal problem, and the reduced master serves to restart the procedure should any element of the decomposed primal be infeasible.

In order to generate a set of solvent blend candidates, the entire problem must be solved repeatedly, adding integer molecular cuts to eliminate previous optimal solutions.

Solvent Verification: Process Level

Having generated a set of potential agent molecule blends, a verification step is required to: (i) test the validity of the generated blends and screen out candidates which are suboptimal or even infeasible due to equilibrium nonconvexities or inaccurate property prediction; (ii) determine if the blends are able to prevent pollution on a process wide level; (iii) accurately estimate the operating cost of the process, as in many cases, blends with excellent environmental performance might be prohibitively expensive. Given details of the manufacturing technology and the global emissions inventory of a facility operating at a fixed production rate, the full plant-wide solvent blend selection problem may be expressed as follows [**P**]

$$\min \left[\textbf{Economic Criterion, Environmental Impact}\right] \tag{88}$$

s.t.

Material and Energy Balances
Feasibility Constraints
Design Equations and Process Specs
Alternative Solvents and Topologies
Input-Output Relationships for all Associated Processes

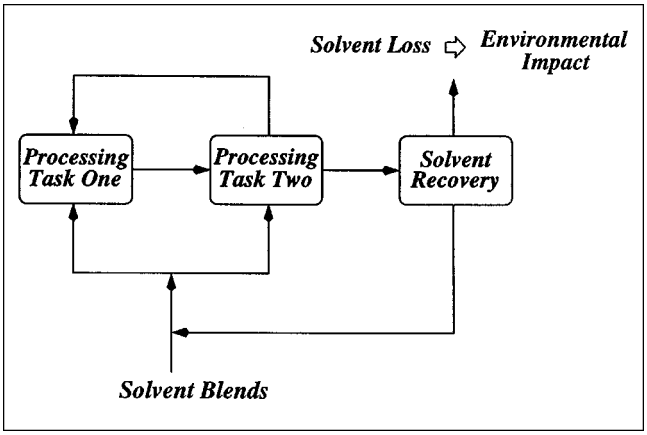

Figure 4. Plant-wide solvent blend design with optimal reuse, recycling, and recovery.

The aim is to minimize global environmental impact (or cost) by simultaneously accounting for solvent re-use, recycling, and recovery (see Figure 4). The full model comprises: detailed design equations and sizing constraints, material and energy balances, logical constraints (for unit and material alternatives), explicit economic criteria (cost of solvents, process capital and operating costs), and (global) environmental impact assessment metrics. Together, these elements form a large multiobjective mixed integer, nonlinear optimization problem. For simplicity, in the examples that follow, we perform a short cut verification exercise, as described below.

Examples

Example 1: gas treatment system of an acrylic fiber plant

In the second section, we determined that the objective in our motivating example should be to reduce both the environmental impact and operating cost of the gas treatment system through solvent substitution in column C-1702. We identified two options: (i) to completely replace the solvent water with a new solvent or solvent blend; or (ii) to introduce additional solvent materials into the solvent water. In either case, the new solvent material(s) should be selective towards both ACN *and* DMF and satisfy all property, process, and environmental constraints.

The feed conditions of the contaminated gas fed to column C-1702, and the solvent specifications to be obeyed are listed in Table 6. The following set of candidate structural groups were employed: CH_3, CH_2, CH, C, OH, H_2O, $COCH_3$, $COCH_2$, CHO, CH_3COO, CH_2COO, CH_3O, and CH_2O.

The number of solvent components is limited to a maximum of two, which is sufficient to illustrate our approach. We imposed tight restrictions on the bubble point and solubility parameters of our solvent blend. We bounded the solvent blend bubble point above the maximum boiling point of the nonsolvent molecules to facilitate regeneration with an upper bound in place in order to limit regeneration energy requirements. The solubility parameter bounds are in place to ensure a good solvent/nonsolvent match. In addition, we have the following end-of-pipe limits for the gas effluent

$$\text{Concentration of DMF} < 20 \text{ ppm}$$
$$\text{Concentration of ACN} < 4 \text{ ppm}$$

Clearly, we could have imposed many more pure component and mixture property constraints, but these are sufficient for this illustrative example.

Focusing on point source pollution, the task based solvent design problem results in the set of solvent blends presented in Table 7. This shows the five best solutions taken from the first ten, compared with the most promising task level solution put forward by Pistikopoulos and Stefanis (1998)—decanol with a *CTAM* of 12.4 tn air/h. The environmental impact is expressed in terms of critical air mass (*CTAM*), since the majority of the process waste is emitted directly to air. Unique molecular connectivities exist for all but one of the groups sets generated—that of the lower composition component of solution C. This group set exhibits two possible connectivities, so that the molecule must be either ethyl propionate or methyl butyrate. For brevity, we assume the molecule to be ethyl propionate for the purposes of verification.

Although the algorithm is flexible to select single solvents, all of the solutions are blends. Due to the introduction of tight property constraints, none of the molecules featured here satisfy all of the imposed constraints on their own, nor do any of the single solvents put forward previously by Pistikopoulos and Stefanis (1998). Clearly and as expected, the blends achieve a wider range and combination of properties

Table 6. Feed Stream and Solvent Specifications for C-1702

Solvent Specs	Feed Conditions
No. of solvents ≤ 2	$P = 1.2$ atm
$2 \leq$ No. of groups ≤ 9	$T = 20$°C
Solvent loss ≤ 0.1	ACN = 8 kg/h
Solvent capacity ≥ 0.1	DMF = 32 kg/h
$440 \text{ K} \leq T_s^{b,\text{mix}} \leq 480 \text{ K}$	Air = 78,100 kg/h
$15.4 \leq \delta_{\text{mix}}^{D} \leq 18.4$	
$5 \leq \delta_{\text{mix}}^{P} \leq 20$	
$6 \leq \delta_{\text{mix}}^{H} \leq 12$	

Table 7. Solvent Blend Identification Results for Example 1

	Solvent Blend (mol. %)	CTAM (kg air/h)
A	89.4% Ethylene glycol methyl ether acetate 10.6% Isopropyl acetate	19.97
B	89.2% Ethylene glycol methyl ether acetate 10.8% *N*-Propyl acetate	9.52
C	89.2% Ethylene glycol methyl ether acetate 10.8% Ethyl propionate/methyl butyrate	9.52
D	87.9% Isopropyl acetate 12.1% Di-methyl succinate	124.68
E	88.2% *N*-Propyl acetate 11.8% Di-methyl succinate	41.72
	Decanol (Pistikopoulos and Stefanis, 1998)	12,400

than the individual molecules and are therefore better able to meet the tight property constraints.

The idea of introducing additional materials into the solvent water was immediately rejected since water fails the azeotropy test with ACN. As shown in Figure 5, the effective vapor pressures of water and ACN cross at a composition of approximately 20 mol. % water, yielding an azeotrope at this point.

Thus, all of the solutions involve completely new materials. For each solution, the predicted temperature change in the absorber was less than 2.34 K, validating the isothermal assumption (see the subsection on solution procedure).

The blend compositions and flow rates are determined by the environmental impact objective according to a trade-off between absorption capacity and solvent loss. However, since the end-of-pipe limits for ACN and DMF emissions are so restrictive, the solvent loss is almost totally responsible for the impact at the task level. Thus, the blend compositions and flow rates are selected so as to minimize the environmental impact of solvent loss.

The mole fraction of each solvent in the absorber vapor outlet stream is related to the pure component solvent inlet mole fraction through the equilibrium relationship

$$X_s^{v_{\text{out}}} \le \frac{X_s^{l_{\text{in}}} \gamma_s^{l_{\text{in}}} P_s^{\text{vap}}}{\phi_s^{v_{\text{out}}} P_{\text{tot}}}$$

from Eqs. 62 and 67 for $\mathcal{G}_1 \ge 0$. The *impact* of the loss of each solvent in the vapor outlet is then

$$\text{IMPACT}_s \le \frac{X_s^{l_{\text{in}}} \gamma_s^{l_{\text{in}}} P_s^{\text{vap}} F^{v_{\text{out}}}}{\phi_s^{v_{\text{out}}} P_{\text{tot}} SLV_s}$$

P_{tot}, P_s^{vap}, and SLV_s are fixed, and we assume $\phi_s^{v_{\text{out}}}$ is unity. Since the solvents are similar, in our example $\gamma_s^{l_{\text{in}}}$ varies little with composition and since the solvent loss is small compared to the total gas flow, $F^{v_{\text{out}}}$ varies little. Thus, effectively, the impact of the loss of each solvent is proportional to the solvent mole fraction at the liquid inlet $X_s^{l_{\text{in}}}$ multiplied by a constant which reflects volatility (through P_s^{vap}) and toxicity (through SLV_s).

In our examples, we have only two solvents at a time and the more volatile solvent components were also always the more toxic. Therefore, minimizing the environmental impact of solvent loss corresponded in all cases to minimizing the liquid inlet mole fraction of the more environmentally damaging solvent material, subject to the blend composition bounds which arise from the pure component property-based mixture property constraints. For example, in solution D (Table 7) isopropyl acetate has a *higher* toxicity than dimethyl succinate so the proportion of isopropyl acetate is minimized. However, according to the blend composition bounds, dimethyl succinate has such a high boiling point that only 12.1 mol. % of it is allowed. Thus, the fraction of isopropyl acetate remains high, accounting for the relatively high impact. All the blend compositions are determined by impact/property trade-offs of this kind.

The solvent flows are determined by the interplay between the environmental impact and the ACN and DMF emission limits. In order to minimize solvent loss, the solvent flows are minimized to such an extent that absorption is just sufficient to meet the emission limits.

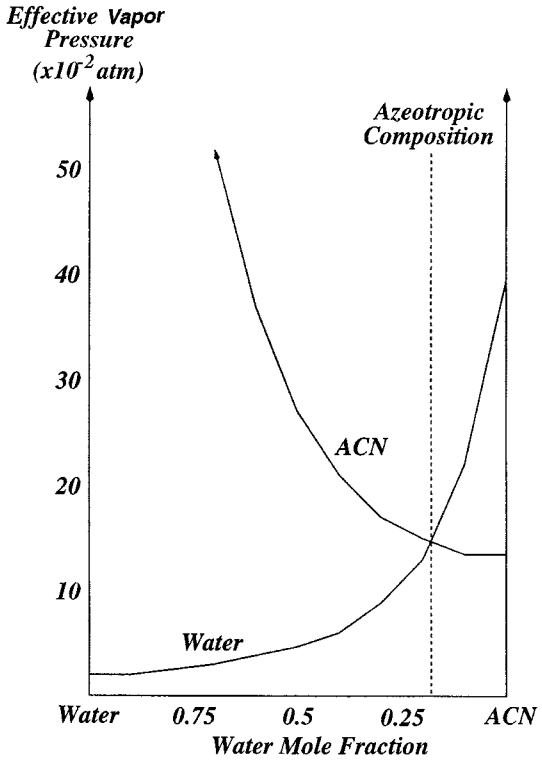

Figure 5. Effective vapor pressures for the water/ACN binary mixture.

The results compare very favorably with both the base case (CTAM 1170 tn air/h) and decanol. The reason for the improvement is the selection of solvents with significantly higher LC50s (that is, lower toxicities). Our solvent molecules have much lower toxicities than those previously put forward, and so exhibit a much improved environmental performance, even though the solvent losses are comparable. Furthermore, all solvent flows are significantly lower than the base case flow, option D has the highest flow with approximately 45,000 kg/h, compared with 200,000 kg/h in the base case.

The blend composition splits are similar for all the blends. This occurs because the same solvent molecules occur in different blends (since the integer cuts written after each solution exclude combinations of solvents and not the individual solvents themselves), and because several of the solvent molecules are very similar (*n*-propyl acetate, isopropyl acetate, and ethyl propionate/methyl butyrate are structural isomers).

Any change in the blend compositions or flow rates will lead to increased impact, but may also lead to violation of the property constraints or the end-of-pipe emission limits for ACN and DMF. Therefore, it is very important to look at the sensitivities of the solutions. The sensitivities of solution A with respect to changes in solvent blend composition at fixed (minimum) flow rate, and with respect to changes in solvent blend flow rate at fixed composition, are shown in Figures 6 and 7, respectively.

As explained above, each solution lies at a composition bound so that certain small changes in the solvent composi-

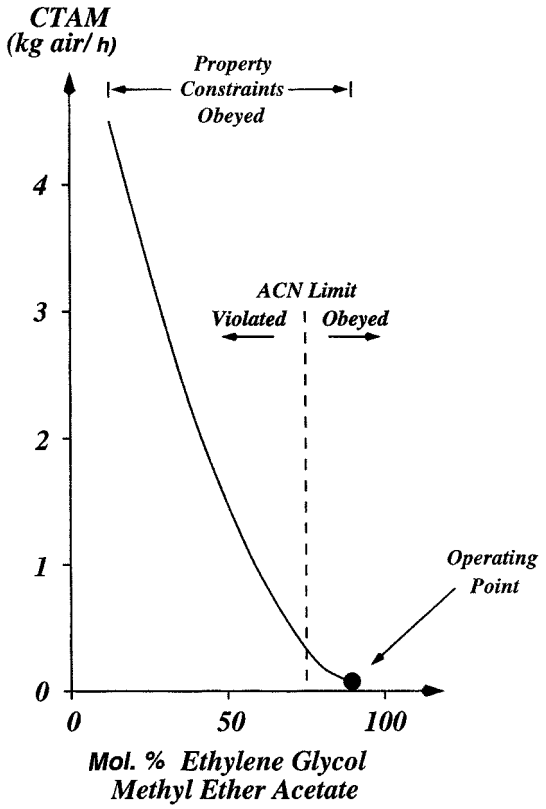

Figure 6. Impact sensitivity to composition: Solution A: Example 1.

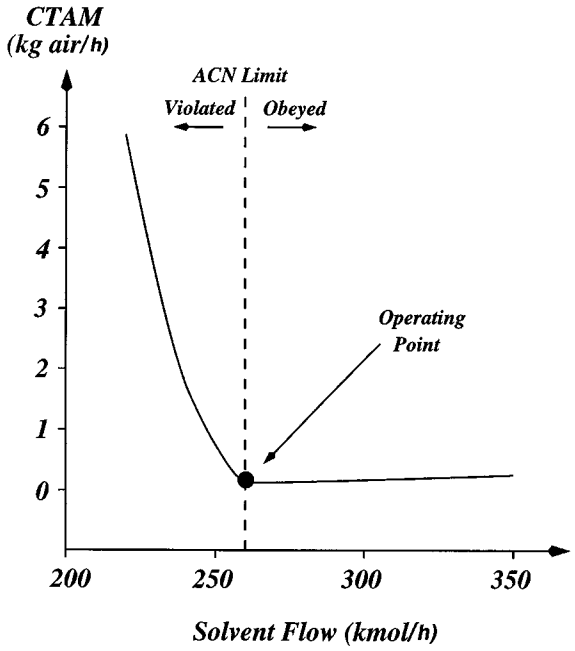

Figure 7. Impact sensitivity to flow: Solution A: Example 1.

tion may lead to property constraint violation. Since the composition of the solvent blend also affects the equilibrium at the bottom of the column (through the activity coefficients) and the solvent loss, changes in composition also affect the impact and may lead to violation of the ACN and DMF emission limits (see Figure 6).

Any reduction in solvent flow immediately leads to violation of the ACN or DMF emission limits, and a corresponding large increase in impact. Although an increase in the solvent flow also leads to increased impact, this increase arises from greater solvent loss and is much less severe (Figure 7).

Thus, a realistic operating regime may involve backing-off from the solvent composition and flow rate limits to provide operational flexibility. This may be achieved by increasing the solvent flow rate and altering the blend composition.

Whereas the solvent flow effectively has no upper limit, the blend composition is bounded by the property constraints. Solutions D and E have very tight composition bounds and so are the least flexible options (such as for option D, $0.1187 \leq X^{\text{in}}_{\text{solv1}} \leq 0.1205$). Solutions A through C have a much greater degree of flexibility (such as solution C with $0.1084 \leq X^{\text{in}}_{\text{solv1}} \leq 0.6902$ is the most tightly restricted of these three). This may be an important criterion in the selection of promising solutions.

Taking all of the above into account, solutions B and C are the most promising at the task level.

At the verification level, we consider the overall gas treatment system including regeneration as shown in Figure 8. We consider only this one topology, and we investigate the performance of each solvent blend alternative in turn with a minimum cost objective. In this way, we reduce the verification model to a manageable NLP problem. Sizing models and cost correlations were obtained from Douglas (1988), with additional cost information taken from Suppliers (1998) and environmental impact data from Pistikopoulos and Stefanis (1998) and Gao et al. (1992). Since the solvent blends did not exhibit large deviations from nonideal behavior in the task level calculations, no further mixture property tests were carried out as part of the verification exercise.

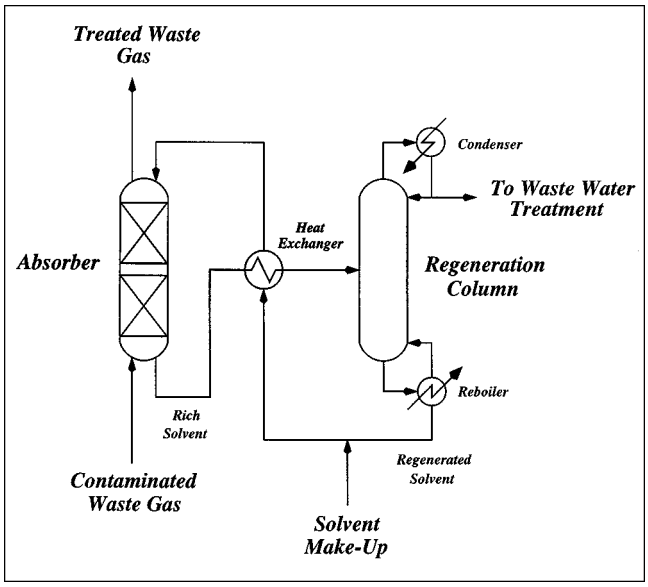

Figure 8. Acrylic fiber plant gas treatment scheme.

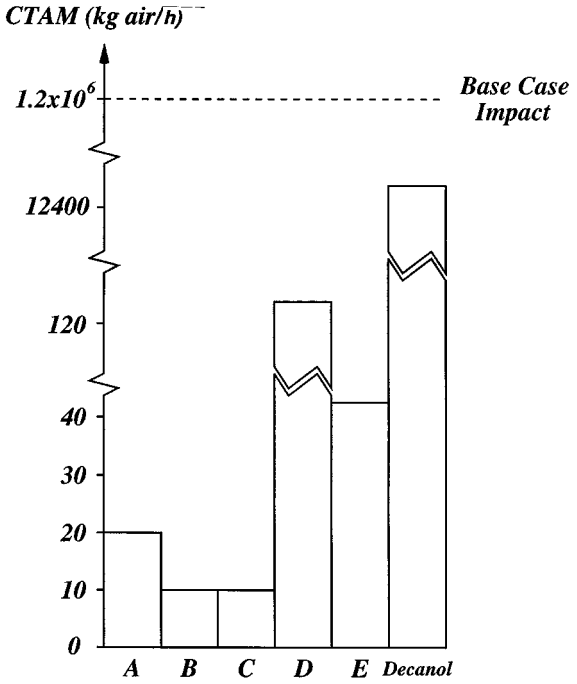

Figure 9. Process CTAM for various solvents: Example 1.

The waste stream from the regeneration column is sent to waste water treatment with no extra environmental penalty. The process environmental and economic results are shown

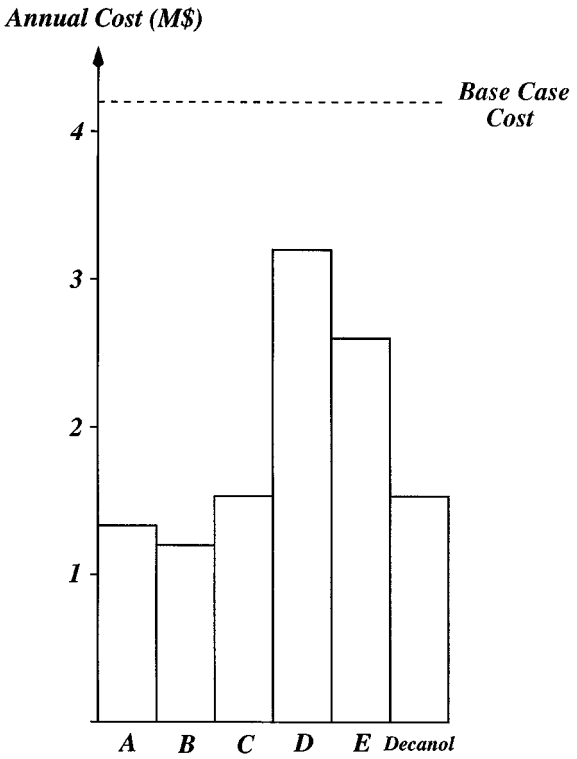

Figure 10. Annual treatment cost for various solvents: Example 1.

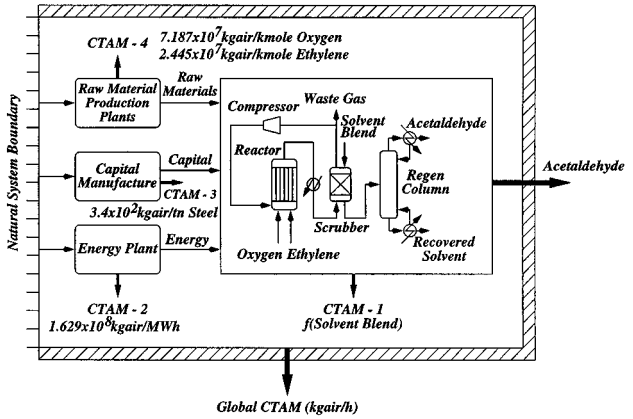

Figure 11. Acetaldehyde production process.

against the base case and decanol in Figures 9 and 10, respectively. In order to make meaningful comparisons, all costings have been calculated using up-to-date cost information taken from *Chemical Marketing Reporter* (1998).

All blend options exhibit significantly improved environmental and economic performance compared with the base case. In environmental impact terms, all blends also outperform decanol. However, in cost terms, decanol remains a competitive alternative. The costs are dominated by the cost of replacement solvent to make up for solvent loss. Options D and E perform worse than decanol, since these blends contain high proportions of volatile components, and, therefore, suffer greater solvent loss. However, options A, B, and C suffer much less solvent loss, and so outperform decanol.

Option B represents the most promising alternative at the process level with the lowest impact and the lowest costs of all due to modest solvent loss and regeneration energy requirements. However, since the performance of options A, B, and C is similar, all three options should be considered further.

Example 2: acetaldehyde recovery

This example concerns the acetaldehyde production process shown in Figure 11. Our objective is to recover acetaldehyde (CH_3CHO) from the reactor effluent so that unreacted oxygen and ethylene may be recycled. The same set of candidate structural groups is used as before.

The feed conditions of the gas to the scrubber and the solvent design specifications are presented in Table 8. Again,

Table 8. Scrubber Feed Stream and Solvent Specifications, Acetaldehyde Example

Solvent Specs	Feed Conditions
No. of solvents ≤ 2	$P = 2$ atm
$2 \leq$ No. of groups ≤ 9	$T = 37$°C
Solvent loss ≤ 0.05	146.5 kmol/h
Solvent capacity ≥ 0.1	23 mol. % ethylene
$T_s^{b,\text{mix}} \geq 310$ K	68 mol. % acetaldehyde
$14.5 \leq \delta_{\text{mix}}^{D} \leq 17.5$	9 mol. % oxygen
$7 \leq \delta_{\text{mix}}^{P} \leq 12$	
$9 \leq \delta_{\text{mix}}^{H} \leq 14$	

Table 9. Solvent Blend Identification Results for Example 2

	Solvent Blend (mol. %)	CTAM (kg air/h)
A	45.5% Methyl acetate 54.5% Ethylene glycol methyl ether	8.87
B	16.7% Methyl propyl ketone 83.3% Ethylene glycol methyl ether	31.66
C	32.9% Methyl ethyl ketone 67.1% Propanol	123.92
Methanol (Pistikopoulos and Stefanis, 1998)		82
Butanal		3197.83

the number of solvents is limited to two, and we have imposed tight restrictions on the solvent solubility parameters in order to ensure a good solvent/nonsolvent match. In this case, the solvent should be selective towards acetaldehyde only, and the solubility parameter bounds are set accordingly. Since the scrubber effluent is recycled to the reactor, the solvent loss specification is tightened.

Focusing on point source pollution, the task-based solvent design problem results in the set of solvent blends shown in Table 9. This shows the three best solutions taken from the first fifteen, all of which were blends. Unique molecular connectivities exist for all the group sets generated.

Also shown in Table 9 is the most promising solution put forward by Pistikopoulos and Stefanis (1998)—methanol with a CTAM of 82 kg/h and butanal. Butanal was the *only* single molecule found to obey all solvent *property* specifications. However, it violates the solvent loss constraint and has a high toxicity, and so exhibits a correspondingly poor environmental performance. Methanol also violates the solvent loss constraint, but is much less toxic. The critical air mass is again used to express the environmental impact since the majority of the task (and global) process wastes are emitted directly to air.

For each blend shown, the predicted temperature change in the scrubber was less than 15 K, validating the isothermal assumption (see the subsection on solution procedure).

The solvent blend compositions are determined by impact/property trade-offs as before. However, there is no limit on the acetaldehyde mole fraction in the waste gas in this case. Thus, the solvent blend flow rates are determined directly by the impact objective according to the trade-off between absorption capacity and solvent loss. Since acetaldehyde is very much more toxic than any of the solvents featured in the blends, the quantity of acetaldehyde in the scrubber exit gas is very small. Thus, as before, the task level impact is dominated by solvent loss, and the solvent flows are set accordingly. All three blend solutions exhibit sufficient flexibility in both solvent flow and composition. Solution A is the most tightly bounded in terms of composition, with a 15 mol. % range.

Options A and B compare very favorably with methanol due to a combination of low solvent loss and toxicity. Option C compares less well. Although it suffers less solvent loss than methanol, both methyl ethyl ketone and propanol are more toxic than methanol, so that the overall impact is greater. Thus, options A and B represent the most promising alternatives at the task level. Butanal is rejected at this stage.

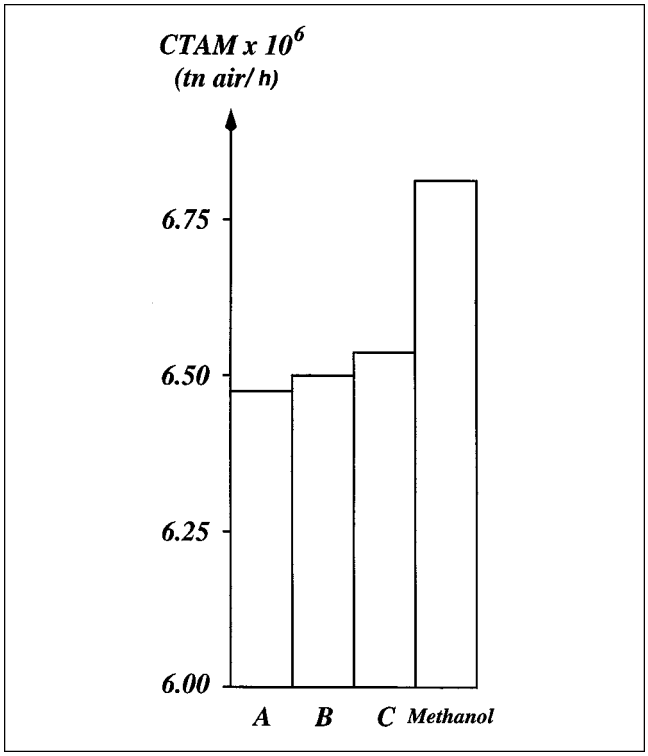

Figure 12. Global process CTAM for various solvents: Example 2.

At the verification level, we consider the global acetaldehyde production process shown in Figure 11. As before, we consider only one process topology and each solvent blend is investigated in turn with a minimum cost objective. The global process environmental and economic results are shown against methanol in Figures 12 and 13, respectively. As be-

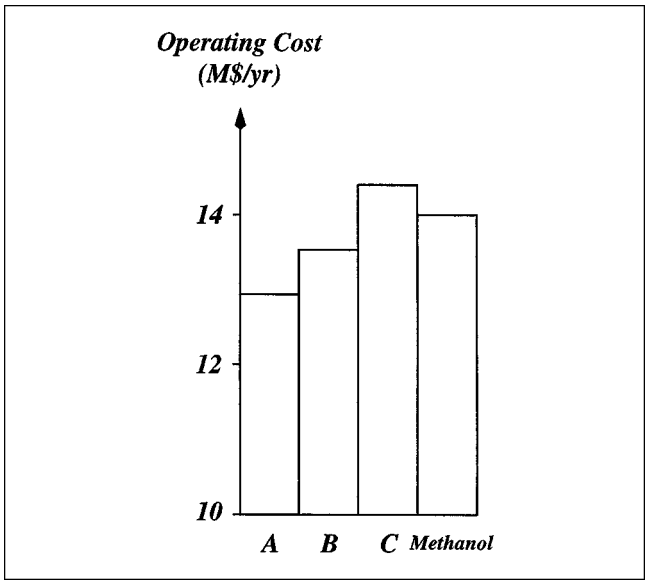

Figure 13. Global process operating cost for various solvents: Example 2.

fore, no further mixture property calculations were performed.

Since very little ethylene or oxygen is absorbed in each case, the raw material demands associated with each solvent option are very similar. Thus, the input wastes associated with raw material production are very similar for each alternative. However, the wastes associated with energy generation vary according to the difficulty of the regeneration separation and are quite different in each case. Since these wastes are very much larger than the task level impacts, the energy wastes dominate the global process impact figures.

Methanol has the lowest boiling point and highest volatility of all the solvent alternatives, and, therefore, the most difficult regeneration separation. Thus, the regeneration column, reflux flow, and energy demand associated with methanol are the largest, so that methanol suffers the highest global process impact figure.

All blend options outperform methanol in terms of impact. Option A exhibits the easiest regeneration separation with the lowest energy demand, and so represents the most promising solution in terms of global process impact. Options B and C have progressively more difficult separations, and so give progressively poorer global impact results.

The cost of replacement solvent to make up for solvent loss dominates the operating cost results. Although methanol suffers the largest solvent loss and the worst global impact, it is the cheapest solvent and so remains a competitive alternative in cost terms. However, since methanol violates the solvent loss constraint, a vapor recovery system would be required to remove methanol from the recycle stream, which would increase the operating cost and make methanol a less competitive alternative.

In any case, blend options A and B outperform methanol in cost terms with option A returning the most attractive results, due to a combination of low solvent loss and low solvent cost.

Thus, option A represents the most promising alternative at the global process level, with the lowest global impact and operating cost. However, options B and C are not far behind, so that all three alternatives are worthy of further investigation.

Conclusions

A systematic procedure to select optimal solvent blend alternatives for nonreactive, multicomponent gas absorption processes has been presented in this article. The approach is based on the solvent design technique proposed by Pistikopoulos and Stefanis (1998). The three step framework involves the identification of agent-based process operations, determining a list of solvent candidates based on explicit environmental constraints and verification of their performance on a plant-wide basis through the application of the principles of MEIM.

Through the introduction of new property and operational feasibility constraints, a refined task-based problem formulation, and a new solution procedure, we have been able to develop very promising solvent blend alternatives for both single and multicomponent absorption tasks involving tight property constraints. Through two example problems, we have shown that solvent blends are better able to meet these constraints than single solvents while simultaneously offering more environmentally *and* economically promising solutions.

Notation

$a_{nj,mj}$ = interaction parameter from VLE data
l_i = molecular parameter
$n_{i,j}$ = number of groups j in molecule i
Q_i = Van der Waals area of component i
q_j = Van der Waals area of group j
r_j = Van der Waals volume of group j
R_i = Van der Waals volume of component i
T_{oper} = operating temperature
X_i^l = exit mole fraction of component i in the liquid mixture
X_j = mole fraction of group j in the mixture
$XC_{i,j}$ = group mole fraction in the reference solution
z = lattice coordination number, usually taken as 10
Θ_j = group Van der Waals area fraction in the mixture
$\Theta C_{i,j}$ = group Van der Waals area fraction in the reference solution
θ_i = Van der Waals area fraction of component i in the mixture
ϕ_i = Van der Waals volume fraction of component i in the mixture
$\psi_{nj,mj}$ = group interaction parameter
Γ_j = residual activity coefficient of group j in the mixture
$\Gamma_{i,j}^o$ = residual activity coefficient of group j in a reference solution containing only molecules of type i

Literature Cited

Acevedo, J., and E. N. Pistikopoulos, ''A Parametric MINLP Approach for Process Synthesis Problems under Uncertainty,'' *Ind. and Eng. Chemistry Res.*, **35**, 147 (1996).

Archer, W. L., *Industrial Solvents Handbook*, Dekker, New York (1996).

Barton, A. F. M., *CRC Handbook of Solubility Parameters and Other Cohessive Parameters*, CRC Press, Boca Raton, FL (1983).

Bollis, G., L. PiDace, and F. Fabrocini, ''A Machine Learning Tool for Computer Aided Molecular Design,'' *Int. Conf. on Tools for Artificial Intelligence*, San Jose, CA (1991).

Brignole, E. A., S. Bottini, and R. Gani, ''A Strategy for the Design and Selection of Solvents for Separation Processes,'' *Fluid Phase Equilib.*, **29**, 125 (1986).

''Chemical Suppliers, Chemical Prices,'' *Chem. Marketing Reporter*, **253**, 25 (1998).

Chen, F., J. Holten-Andersen, and H. Tyle, ''New Developments of the UNIFAC Model for Environmental Application,'' *Chemosphere*, **26**, 1325 (1993).

Churi, N., and L. E. K. Achenie, ''Novel Mathematical Programming Model for Computer Aided Molecular Design,'' *Ind. and Eng. Chemistry Res.*, **35**, 3788 (1996).

Constantinou, L., C. Jaksland, K. Bagherpour, and R. Gani, ''Application of the Group Contribution Approach to Tackle Environmentally Related Problems,'' *Pollution Prevention through Process and Product Modifications*, AIChE Symp. Series, **303**, 105 (1994).

Constantinou, L., K. Bagherpour, R. Gani, J. A. Klein, and D. T. Wu, ''Computer Aided Product Design: Problem Formulations, Methodology, and Applications,'' *Computers Chem. Eng.*, **20**, 685 (1996).

Derringer, G. C., and R. L. Markham, ''A Computer-Based Methodology for Matching Polymer Structures with Required Properties,'' *J. of Appl. Poly. Sci.*, **30**, 4609 (1985).

Douglas, J. M., *Conceptual Design of Chemical Processes*, 1st ed., McGraw-Hill, New York (1988).

Dunn, R. F., A. M. Dobson, and M. M. El-Halwagi, *Optimal Design of Environmentally-Acceptable Solvent Blends for Coatings*, in press (1997).

Duvedi, A., and L. E. K. Achenie, ''On the Design of Environmentally Benign Refrigerant Mixtures: A Mathematical Programming Approach,'' *Computers Chem. Eng.*, **21**, 915 (1996a).

Duvedi, A. P., and L. K. Achenie, ''Designing Environmentally Safe Refrigerants using Mathematical Programming,'' *Chem. Eng. Sci.*, **51**, 3727 (1996b).

Fredenslund, A., R. L. Jones, and J. M. Prausnitz, "Group Contribution Estimation of Activity Coefficients in Nonideal Liquid Mixtures," *AIChE J.*, **21**, 1086 (1975).
Gani, R., and A. Fredenslund, "Computer Aided Molecular and Mixture Design with Specified Property Constraints," *Fluid Phase Equilib.*, **82**, 39 (1993).
Gani, R., and E. A. Brignole, "Molecular Design of Solvents for Liquid Extraction Based on UNIFAC," *Fluid Phase Equilib.*, **13**, 331 (1983).
Gani, R., B. Neilsen, and A. Fredenslund, "Group Contribution Approach to Computer-Aided Molecular Design," *AIChE J.*, **37**, 1318 (1991).
Gani, R., N. Tzouvaras, P. Rasmussen, and A. Fredenslund, "Prediction of Gas Solubility and Vapor-Liquid Equilibria by Group Contribution," *Fluid Phase Equilib.*, **47**, 105 (1989).
Gao, C., R. Govind, and H. Tabak, "Application of the Group Contribution Method for Predicting the Toxicity of Organic Chemicals," *Environ. Toxicol. and Chemistry*, **11**, 631 (1992).
Geoffrion, A. M., "Generalized Benders Decomposition," *J. Optimization Theory Applications*, **10**, 237 (1974).
Godfrey, J. C., and M. J. Slater, *Liquid-Liquid Extraction Equipment*, Wiley, Chichester, U.K. (1994).
Gordeeva, E. V., M. S. Molcharova, and N. S. Zefirov, "General Methodology and Computer Program for the Exhaustive Restoring of Chemical Structures by Molecular Connectivity Indices. Solution of the Inverse Problem in QSAP/QSPR," *Tetrahedron Comput. Methodol.*, **3**, 389 (1990).
Hankinson, R. W., and G. H. Thomas, "A New Correlation for Saturated Densities of Liquids and Their Mixtures," *AIChE J.*, **25**, 653 (1979).
Hansen, C. M., "The Three Dimensional Solubility Parameter—Key to Print Component Affinities: 1. Solvents, Plasticizers, Polymers and Resins," *J. Paint. Tech.*, **39**, 104 (1967).
Harding, S. T., C. D. Maranas, C. M. McDonald, and C. A. Floudas, "Locating all Homogeneous Azeotropes in Multicomponent Mixtures," *Ind. Eng. Chem. Res.*, **36**, 160 (1997).
Hildebrand, J. H., and R. L. Scott, *The Solubility of Nonelectrolytes*, Reinhold Publishing Co., New York (1950).
Horvath, A. L., *Molecular Design*, 1st ed., Elsevier, London (1992).
Jaksland, C., and R. Gani, "An Integrated Approach to Process/Product Design and Synthesis Based on Properties-Process Relationship," *Comput. Chem. Eng.*, **20**, S151 (1996).
Joback, K. G., "Unified Approach to Physical Property Estimation using Multivariate Statistical Techniques," BS Thesis, MIT, Cambridge, MA (1984).
Joback, K. G., "Designing Molecules Possessing Desired Physical Property Values," PhD Thesis, MIT, Cambridge, MA (1989).
Joback, K. G., and G. Stephanopoulos, "Designing Molecules Possessing Desired Physical Property Values," *Proc. FOCAPD*, CACHE Corp., Austin, TX, **11**, 631 (1989).
Joback, K. G., and G. Stephanopoulos, "Designing Molecules Possessing Desired Physical Property Values," *Adv. in Chem. Eng.*, **21**, Intelligent Systems in Process Engineering, Academic Press, London (1995).
Kier, L. B., H. H. Lowell, and J. F. Frazer, "Design of Molecules from Quantitative Structure-Activity Relationship Models: 1. Information Transfer between Path and Vertex Degree Counts," *J. Chem. Inf. Comput. Sci.*, **33**, 142 (1993).
Klein, J. A., D. T. Wu, and R. Gani, "Computer Mixture Design with Specified Property Constraints," *Comput. Chem. Eng.*, **16**, S229 (1992).
Lo, T. C., M. H. I. Baird, and C. Hanson, *Handbook of Solvent Extraction*, Wiley, New York (1983).
Macchietto, S., O. Odele, and O. Omatsone, "Design of Optimal Solvents for Liquid-Liquid Extraction and Gas Absorption Processes," *Trans. IChemE*, **68**, 429 (1990).
Magnussen, T., "UNIFAC Parameter Table for Prediction of Liquid-Liquid Equilibria," *Ind. Eng. Chem. Process Des. Dev.*, **20**, 331 (1981).
Maranas, C. D., "Optimal Computer-Aided Molecular Design: A Polymer Design Case Study," *Ind. Chem. Eng. Res.*, **35**, 3403 (1996).
Maranas, C. D., "Optimal Molecular Design Under Property Prediction Uncertainty," *AIChE J.*, **43**, 1250 (1997).
Mavrovouniotis, M. L., "Estimation of Properties from Conjugate Forms of Molecular Structures: The ABC Approach," *Ind. Eng. Chem. Res.*, **23**, 1943 (1990).
McDonald, C. M., and C. A. Floudas, "Global Optimization and Analysis for the Gibbs Free Energy Function Using the UNIFAC, Wilson and ASOG Models," *Ind. Eng. Chem. Res.*, **34**, 1674 (1995).
Modi, A., J. P. Aumond, M. Mavrovouniotis, and G. Stephanopoulos, "Rapid Plant-Wide Screening of Solvents for Batch Processes," *Comput. Chem. Eng.*, **20**, S375 (1996).
Naser, S. F., and R. L. Fournier, "A System for the Design of an Optimum Liquid-Liquid Extractant Molecule," *Comput. Chem. Eng.*, **15**, 397 (1991).
Odele, O., and S. Macchietto, "Computer Aided Molecular Design: A Novel Method for Optimal Solvent Selection," *Fluid Phase Equilibria*, **82**, 47 (1993).
Papalexandri, K. P., and E. N. Pistikopoulos, "Generalized Modular Representation Framework for Process Synthesis," *AIChE J.*, **42**, 1010 (1996).
Pistikopoulos, E. N., and S. K. Stefanis, "Optimal Solvent Design for Environmental Impact Minimization," *Comput. Chem. Eng.*, **6**, 717 (1998).
Pistikopoulos, E. N., S. K. Stefanis, and A. G. Livingston, "A Methodology for Minimum Environmental Impact Analysis," *Pollution Prevention through Process and Product Modifications*, AIChE Symp. Ser., **90**, 139 (1994).
Pretel, E. J., P. A. Lopez, S. B. Bottini, and E. A. Brignole, "Computer-Aided Molecular Design of Solvents for Separation Processes," *AIChE J.*, **40**, 1349 (1994).
Rahim Ismail, S., K. P. Papalexandri, and E. N. Pistikopoulos, "Separation of Non-Ideal Mixtures Based on Mass/Heat Exchange Principles. The Entrainer Selection and Sequencing Problem," *Computers Chem. Eng.*, **21**, S211 (1997).
Reid, R. C., J. M. Prausnitz, and B. E. Poling, *The Properties of Gases and Liquids*, McGraw-Hill, New York (1987).
Small, P. A., "Some Factors Affecting the Solubility of Polymers," *J. of Appl. Chemistry*, **3**, 71 (1953).
Smith, J. M., H. C. Van Ness, and M. M. Abbott, *Introduction to Chemical Engineering Thermodynamics*, 5th ed., McGraw-Hill, New York, p. 501 (1996).
Stefanis, S. K., and E. N. Pistikopoulos, "A Methodology for Environmental Risk Assessment for Industrial Non-Routine Releases," *Ind. Eng. Chem. Res.*, **36**, 3694 (1997).
Stephanopoulos, G., and D. W. Townsend, "Synthesis in Process Development," *Chem. Eng. Res. and Dev.*, **64**, 160 (1986).
Thompson, G. H., K. R. Brobst, and R. W. Hankinson, "An Improved Correlation for Densities of Compressed Liquids and Liquid Mixtures," *AIChE J.*, **28**, 671 (1982).
Vaidyanathan, R., and M. El-Halwagi, "Computer-Aided Synthesis of Polymers and Blends with Target Properties," *Ind. Eng. Chem. Res.*, **35**, 627 (1996).
Venkatasubramanian, V., K. Chan, and J. M. Caruthers, "Computer-Aided Molecular Design Using Genetic Algorithms," *Computers Chem. Eng.*, **18**, 833 (1994).

Appendix A: Chemical Feasibility Constraints

Archer (1996) provides a list of 289 industrial solvents. Of these, 96 compounds were rejected (see Appendix B), the remaining 193 contributed to the development of the following set of chemical feasibility constraints for solvent molecules designed using the group set employed in our example problems (see the Examples section).

The maximum and minimum numbers of groups are limited as follows

$$\sum_j n_{sj} \le 9, \quad \forall s \in S \tag{A1}$$

$$\sum_j n_{sj} \ge 1, \quad \forall s \in S \tag{A2}$$

The maximum number of functional groups is limited to four

$$n_{s,CH_3COO} + n_{s,CH_2COO} + n_{s,CH_3O} + n_{s,CH_2O} + n_{s,COCH_3} + n_{s,COCH_2} + n_{s,CHO} + n_{s,OH} \le 4, \quad \forall s \in S \quad (A3)$$

The maximum occurrence for each functional group is limited as follows

$$n_{s,CH_3COO} \le 1, \quad \forall s \in S \quad (A4)$$

$$n_{s,CH_2COO} \le 2, \quad \forall s \in S \quad (A5)$$

$$n_{s,COCH_3} \le 1, \quad \forall s \in S \quad (A6)$$

$$n_{s,COCH_2} \le 1, \quad \forall s \in S \quad (A7)$$

$$n_{s,CHO} \le 1, \quad \forall s \in S \quad (A8)$$

$$n_{s,OH} \le 3, \quad \forall s \in S \quad (A9)$$

$$n_{s,CH_3O} \le 2, \quad \forall s \in S \quad (A10)$$

$$n_{s,CH_2O} \le 2, \quad \forall s \in S \quad (A11)$$

$$n_{s,ACRY} = 0, \quad \forall s \in S \quad (A12)$$

$$n_{s,DMF} = 0, \quad \forall s \in S \quad (A13)$$

$$n_{s,C} \le 1, \quad \forall s \in S \quad (A14)$$

$$n_{s,CH} \le 1, \quad \forall s \in S \quad (A15)$$

Certain combinations of groups never appear together

$$\sum_t \left(INT_{s,CH_3COO,t} + INT_{s,CH_2COO,t} + INT_{s,COCH_3,t} + INT_{s,COCH_2,t} + INT_{s,CHO,t} + INT_{s,OH,t} \right) \le 1, \quad \forall s \in S \quad (A16)$$

$$\sum_t \left(NT_{s,COCH_3,t} + INT_{s,COCH_2,t} + INT_{s,CH_3O,t} + INT_{s,CH_2O,t} + INT_{s,CHO,t} \right) \le 1, \quad \forall s \in S \quad (A17)$$

$$\sum_t \left(INT_{s,CH_2COO,t} + INT_{s,CH_3O,t} + INT_{s,CH_2O,t} \right) \le 1, \quad \forall s \in S \quad (A18)$$

Certain combinations of groups appear together in at most 2s, 3s, or 4s

$$n_{s,CH_3COO} + n_{s,CH_3O} + n_{s,CH_2O} \le 3, \quad \forall s \in S \quad (A19)$$

$$n_{s,OH} + n_{s,CH_3O} + n_{s,CH_2O} \le 4, \quad \forall s \in S \quad (A20)$$

$$n_{s,CH_3O} + n_{s,CH_2O} \le 2, \quad \forall s \in S \quad (A21)$$

Certain combinations of groups must appear with at least 2 or 3 others

$$\sum_j n_{sj} - \left(n_{s,CH_3COO} + n_{s,CH_3O} \right) \ge \sum_t \left(INT_{s,CH_3COO,t} + INT_{s,CH_3O,t} \right), \quad \forall s \in S \quad (A22)$$

$$\sum_j n_{sj} - \left(n_{s,OH} + n_{s,CH_3O} \right) \ge \sum_t \left(INT_{s,OH,t} + INT_{s,CH_3O,t} \right), \quad \forall s \in S \quad (A23)$$

$$\sum_j n_{sj} - \left(n_{s,OH} + n_{s,CH_2O} \right) \ge 2 \sum_t \left(INT_{s,OH,t} + INT_{s,CH_2O,t} \right) - 1, \quad \forall s \in S \quad (A24)$$

Water must appear alone if at all

$$n_{s,H_2O} \le 1, \quad \forall s \in S \quad (A25)$$

$$\sum_j n_{sj} + 8\, INT_{s,H_2O,1} \le 9, \quad \forall s \in S \quad (A26)$$

Appendix B: Solubility Parameter Group Contribution Methods

Archer (1996) provides Hansen solubility parameters for some 289 solvents. We reject 96 of these compounds due to: incompatibility with UNIFAC group structures, use of undesirable groups (such as those containing chlorine), insufficient structural information (for some proprietary solvents) or over complicated structure. The remaining 193 compounds may be represented using 33 UNIFAC groups. We have performed both linear and least-squares regressions in order to identify the contributions of the groups towards each of the Hansen solubility parameters. The linear regression yielded the more accurate results, with a lower mean of absolute error magnitudes in each case.

Linear regression formulation

Our regression formulation is as follows

$$\text{minimize} \sum_i \epsilon_i^{mag} \quad (B1)$$

s.t.

$$\sum_j n_{ij} \Delta_{j,\delta^*} + \alpha^* = \delta_i^{*pred}, \quad \forall i \in I \quad (B2)$$

$$\delta_i^{*pred} = \delta_i^{*lit} + \epsilon_i^{abs}, \quad \forall i \in I \quad (B3)$$

$$\epsilon_i^{abs} = \epsilon_i^{+} - \epsilon_i^{-}, \quad \forall i \in I \quad (B4)$$

$$\epsilon_i^{mag} = \epsilon_i^{+} + \epsilon_i^{-}, \quad \forall i \in I \quad (B5)$$

$$\epsilon_i^{mag} \ge 0, \quad \epsilon_i^{+} \ge 0, \quad \epsilon_i^{-} \ge 0, \quad \forall i \in I \quad (B6)$$

The form of the equations is the same for each of the three solubility parameters. The asterisk represents D, P, or H for the dispersive forces, polar forces, or hydrogen bonding solubility parameters, respectively.

Table B1. Solubility Parameter UNIFAC Group Contributions

Parameter	UNIFAC Group								
	CH_3	CH_2	CH	C	$CH_2=CH$	ACH	AC	$ACCH_3$	
δ^D	−0.7684	−0.0053	0.3947	0.5632	−0.7368	0.2614	0.6430	−0.0386	
δ^P	−0.3531	−0.4750	−0.6719	−0.8312	3.6844	−0.2667	0.6521	−0.6667	
δ^H	−1.5571	−0.8429	−0.1286	−1.1000	−0.2000	−1.2643	1.9786	−1.8214	
	$ACCH_2$	ACCH	OH	ACOH	CH_3CO	CH_2CO	CH_3COO	CH_2COO	
δ^D	0.1509	−0.0061	0.1105	−0.7228	−0.2526	−0.0895	0.0605	0.4579	
δ^P	−0.2385	−0.6458	4.1906	5.6833	5.8812	5.3812	2.7281	2.0812	
δ^H	−0.5786	0.0000	7.9857	12.4357	−0.8286	1.9429	2.1857	2.5143	
	HCOO	CH_3O	CH_2O	CH−O	CH_2NH_2	$CHNH_2$	CH_2NH	CH_2N	
δ^D	−0.5579	−0.9816	−0.2132	−0.6526	−0.1158	−1.0237	0.1211	0.1632	
δ^P	5.8281	1.3812	1.0125	3.3844	2.7000	1.5375	0.1312	4.9656	
δ^H	0.9429	−0.5143	−0.5143	−2.1000	2.5429	−1.3929	0.2000	−2.5000	
	$ACNH_2$	COOH	COO	CH_2CN	$ACNO_2$	CHO	CH_2NO_2	CH_3N	H_2O
δ^D	1.2614	−1.5632	0.4263	−0.7632	3.1991	−0.0605	0.9053	1.3974	−0.3316
δ^P	3.0333	4.9531	1.8625	11.2531	8.4292	5.4406	9.4062	5.2125	20.1
δ^H	6.6643	5.7000	−1.2571	−1.5143	−5.7571	3.7000	−3.3571	−0.7000	4.9429

Our objective is to identify the set of group contribution parameters Δ_{j,δ^*}, which minimizes the sum of the *magnitudes* of the absolute errors between the literature values of the solubility parameters ($\delta_i^{*\text{lit}}$) and the predicted values ($\delta_i^{*\text{pred}}$).

Since the absolute error ϵ_i^{abs} may be positive or negative, we decompose the error into a positive contribution ϵ_i^+ and a negative contribution ϵ_i^-. Both ϵ_i^+ and ϵ_i^- are defined as positive variables so that ϵ_i^- appears with a negative sign in Eq. B4. Thus, if the absolute error is positive, ϵ_i^+ is non-zero and ϵ_i^- is zero, while for a negative absolute error, ϵ_i^+ is zero, but ϵ_i^- is non-zero. We can then write the magnitude of the absolute error ϵ_i^{mag} as the sum of ϵ_i^+ and ϵ_i^-. α^* is a constant to be determined by the regression.

Group parameters

The group contribution parameters for our 33 UNIFAC groups towards the dispersive, polar forces and hydrogen bonding Hansen solubility parameters are given in Table B1.

Prediction equations

Using the above contributions we predict the solubility parameters using the following equations

$$\delta_i^D = \sum_j n_{ij}\Delta_{j,\delta^D} + 16.8316, \quad \forall i \in I$$

$$\delta_i^P = \sum_j n_{ij}\Delta_{j,\delta^P} + 3.4000, \quad \forall i \in I$$

$$\delta_i^H = \sum_j n_{ij}\Delta_{j,\delta^H} + 9.8571, \quad \forall i \in I$$

Table B2. Solubility Parameter Prediction Errors

Parameter	ϵ^{mag}	ϵ_i^{mag}							
		>0.5	>1.0	>1.5	>2	>3	>4	>5	>10
δ_i^D	0.5045	34%	15%	7%	4%	2%	0.5%	0%	0%
δ_i^P	1.2754	57%	39%	22%	18%	14%	7%	6%	1%
δ_i^H	1.5829	69%	56%	43%	32%	17%	8%	3%	0.5%

Prediction errors

The mean of the *magnitudes* of the absolute errors ϵ^{mag} in the predicted solubility parameters for the training set of 193 compounds are given in Table B2. This table also shows the percentage of compounds from the training set which exhibit absolute error magnitudes in the predicted solubility parameters which are above certain limits.

The means of absolute error magnitudes indicate acceptable accuracy for all three prediction techniques. Furthermore, the spread of errors reveals that it is unlikely that we will suffer an absolute prediction error magnitude of more than three of four. This too is acceptable.

Appendix C: Standard UNIFAC Equations

Equations

Overall Activity Coefficient Equation

$$\ln \gamma_i = \ln \gamma_i^c + \ln \gamma_i^r \tag{C1}$$

Combinatorial Part

$$\ln \gamma_i^c = \ln\left(\frac{\phi_i}{X_i^l}\right) + \frac{z}{2} Q_i \ln\left(\frac{\theta_i}{\phi_i}\right) + l_i - \phi_i\left(\frac{\sum_{ii} X_{ii}^l l_{ii}}{X_i^l}\right) \tag{C2}$$

$$\phi_i = \left(\frac{X_i^l R_i}{\sum_{ii} X_{ii}^l R_{ii}}\right) \tag{C3}$$

$$\sum_i \phi_i = 1 \tag{C4}$$

$$\theta_i = \left(\frac{X_i^l Q_i}{\sum_{ii} X_{ii}^l Q_{ii}}\right) \tag{C5}$$

$$\sum_i \theta_i = 1 \tag{C6}$$

$$l_i = \frac{z}{2}(R_i - Q_i) - (R_i - 1) \tag{C7}$$

$$R_i = \sum_j n_{i,j} r_j \tag{C8}$$

$$Q_i = \sum_j n_{i,j} q_j \tag{C9}$$

Residual Part

$$\ln \gamma_i^r = \sum_j n_{ij}\left(\ln \Gamma_j - \ln \Gamma_{ij}^o\right). \tag{C10}$$

Mixture Residual Group Activity Coefficient

$$\ln \Gamma_j = q_j \left\{ 1 - \ln\left(\sum_{mj} \Theta_{mj} \psi_{mj,j} \right) - \sum_{mj} \left(\frac{\Theta_{mj} \psi_{j,mj}}{\sum_{nj} \Theta_{nj} \psi_{nj,mj}} \right) \right\} \tag{C11}$$

$$\Theta_j = \left(\frac{q_j X_j}{\sum_{mj} q_{mj} X_{mj}} \right) \tag{C12}$$

$$\sum_j \Theta_j = 1 \tag{C13}$$

$$X_j = \left(\frac{\sum_i n_{i,j} X_i^l}{\sum_i \sum_{mj} n_{i,mj} X_i^l} \right) \tag{C14}$$

$$\sum_j X_j = 1 \tag{C15}$$

Pure Component Residual Group Activity Coefficient

$$\ln \Gamma_{i,j}^o = q_j \left\{ 1 - \ln\left(\sum_{mj} \Theta C_{i,mj} \psi_{mj,j} \right) - \sum_{mj} \left(\frac{\Theta C_{i,mj} \psi_{j,mj}}{\sum_{nj} \Theta C_{i,nj} \psi_{nj,mj}} \right) \right\} \tag{C16}$$

$$\Theta C_{i,j} = \left(\frac{q_j X C_{i,j}}{\sum_{mj} q_{mj} X C_{i,mj}} \right) \tag{C17}$$

$$\sum_j \Theta C_{i,j} = 1 \tag{C18}$$

$$X C_{i,j} = \left(\frac{n_{i,j}}{\sum_{mj} n_{i,mj}} \right) \tag{C19}$$

$$\sum_j X C_{i,j} = 1 \tag{C20}$$

$$\psi_{nj,mj} = \exp\left(- a_{nj,mj} / T_{\text{oper}} \right) \tag{C21}$$

Appendix D: Reformulation of UNIFAC Equations

Combinatorial part

Substituting for ϕ_i and θ_i in Eq. C2, we obtain the following, noting that X_i^l cancels into ϕ_i (see Eq. C3) in the first and third RHS terms and cancels within the second RHS term

$$\ln \gamma_i^c = \ln\left(\frac{R_i}{\sum_{ii} X_{ii}^l R_{ii}} \right) + \frac{z}{2} Q_i \ln\left(\frac{Q_i \sum_{ii} R_{ii} X_{ii}^l}{R_i \sum_{ii} X_{ii}^l Q_{ii}} \right) + l_i - \left(\frac{R_i \sum_{ii} X_{ii}^l l_{ii}}{\sum_{ii} X_{ii}^l R_{ii}} \right) \tag{D1}$$

Assigning the value of 10 to z in Eq. C7, we obtain the following expression for l_i

$$l_i = 4R_i - 5Q_i + 1 \tag{D2}$$

We now substitute for l_i in the third and fourth RHS terms of Eq. D1

$$l_i - \left(\frac{R_i \sum_{ii} X_{ii}^l l_{ii}}{\sum_{ii} X_{ii}^l R_{ii}} \right) = 4R_i - 5Q_i + 1 - 4\left(\frac{R_i \sum_{ii} X_{ii}^l R_{ii}}{\sum_{ii} X_{ii}^l R_{ii}} \right) + 5\left(\frac{R_i \sum_{ii} X_{ii}^l Q_{ii}}{\sum_{ii} X_{ii}^l R_{ii}} \right) - \left(\frac{R_i \sum_{ii} X_{ii}^l}{\sum_{ii} X_{ii}^l R_{ii}} \right) \tag{D3}$$

We note that the fourth RHS term reduces to $4R_i$, and so cancels with the first RHS term, and, since $\Sigma_i X_i^l = 1$, we can simplify the last RHS term and make a factorization in the last two terms. Substituting this result back into Eq. D1, again assigning the value of 10 to z, we obtain the following

$$\ln \gamma_i^c = \ln\left(\frac{R_i}{\sum_{ii} X_{ii}^l R_{ii}} \right) + 5Q_i \ln\left(\frac{Q_i \sum_{ii} R_{ii} X_{ii}^l}{R_i \sum_{ii} X_{ii}^l Q_{ii}} \right) - 5Q_i + 1 + \left(5 \sum_{ii} X_{ii}^l Q_{ii} - 1 \right)\left(\frac{R_i}{\sum_{ii} X_{ii}^l R_{ii}} \right) \tag{D4}$$

We now expand the logarithms so that Eq. D4 is transformed into Eq. 39. Now, we no longer need to define ϕ_i, θ_i or l_i, so we can eliminate Eqs. C3–C7. This significantly reduces the complexity of the combinatorial part equations and in particular their compositional dependence.

Residual part

Equation C10 may be expanded as follows

$$\ln \gamma_i^r = \sum_j n_{i,j} \ln \Gamma_j - \sum_j n_{i,j} \Gamma_{i,j}^o \tag{D5}$$

The first RHS term represents the sum of the group residual activity coefficient contributions for the groups j in the molecule i in the liquid mixture with all other molecules present (that is, this is the mixture term). The second RHS term represents the sum of the group residual activity coefficient contributions for the groups j in molecule i in a reference solution containing *only* molecule i (that is, this is the pure component term). We now consider these terms separately.

Mixture Terms. Substituting for X_j in Eq. C12, we obtain the following result, noting that the denominators of Eq. C14 cancel

$$\Theta_{mj} = \left(\frac{q_{mj} \sum_i n_{i,mj} X_i^l}{\sum_i Q_i X_i^l} \right) \tag{D6}$$

Substituting this expression into Eq. C11, we now expand the mixture term from Eq. D5, noting that the denominators of Eq. D6 cancel

$$\sum_j n_{i,j} \ln \Gamma_j = \sum_j n_{i,j} q_j - \sum_j n_{i,j} q_j \ln \left(\frac{\sum_{mj} q_{mj} \sum_{ii} n_{ii,mj} X_{ii}^l \psi_{mj,j}}{\sum_{ii} Q_{ii} X_{ii}^l} \right) - \sum_j n_{i,j} q_j \sum_{mj} \left(\frac{q_{mj} \sum_{ii} n_{ii,mj} X_{ii}^l \psi_{j,mj}}{\sum_{nj} q_{nj} \sum_{ii} n_{ii,nj} X_{ii}^l \psi_{nj,mj}} \right) \tag{D7}$$

We extract the denominator of the third RHS term from the internal summation over mj into the external summation over j. This latter operation has the effect of reversing the order of the indices of ψ in the numerator of the third term and of converting $\psi_{nj,mj}$ and $\psi_{nj,j}$ in the denominator (McDonald and Floudas, 1995). We perform this operation so that the numerator and denominator can be evaluated independently for each value of j. Although there are no cancellation opportunities, this simplifies our programming. We also separate the logarithm in the second RHS term.

The resulting equation is as follows, although still complex, we have again reduced the compositional dependence of the system through the cancellations

$$\sum_j n_{i,j} \ln \Gamma_j = Q_i - \sum_j n_{i,j} q_j \ln \left(\sum_{mj} q_{mj} \sum_{ii} n_{ii,mj} X_{ii}^l \psi_{mj,j} \right) + Q_i \ln \left(\sum_{ii} Q_{ii} X_{ii}^l \right) - \sum_j \left(\frac{\sum_{mj} n_{i,mj} q_{mj} q_j \sum_{ii} n_{ii,j} X_{ii}^l \psi_{mj,j}}{\sum_{nj} q_{nj} \sum_{ii} n_{ii,nj} X_{ii}^l \psi_{nj,j}} \right) \tag{D8}$$

Furthermore, we need no longer define Θ_j or X_j so that Eqs. C12–C15 are eliminated. Note also that the argument inside the logarithm of the second RHS term is the same as the denominator of the fourth RHS term. Therefore, we need evaluate this term only once for each value of j.

Pure Component Terms. Substituting for $XC_{i,j}$ in Eq. C17, we obtain the following result, noting that the denominators of Eq. C19 cancel

$$\Theta_{i,mj} = \left(\frac{q_{mj} n_{i,mj}}{Q_i} \right) \tag{D9}$$

Substituting this expression into Eq. C16, we now expand the pure component term from Eq. D5, noting that the denominators of Eq. D9 cancel. As before, we separate the logarithm of the second RHS term and we extract the denominator of the third RHS term from the internal summation over mj into the external summation over j. As before, the indices of ψ are reversed in the numerator and converted from (nj, mj) to (nj, j) in the denominator. The resulting equation is as follows

$$\sum_j n_{i,j} \ln \Gamma_{i,j}^o = \sum_j n_{i,j} q_j - \sum_j n_{i,j} q_j \ln \left(\sum_{mj} q_{mj} n_{i,mj} \psi_{mj,j} \right) + Q_i \ln Q_i - \sum_j \left(\frac{q_j n_{i,j} \sum_{mj} n_{i,mj} q_{mj} \psi_{mj,j}}{\sum_{nj} q_{nj} n_{i,nj} \psi_{nj,j}} \right) \tag{D10}$$

This time we note there is a cancelling opportunity in the fourth RHS term. Since the indices mj and nj run independently, the summations in the numerator and the denominator are equivalent and so cancel out. The term then reduces to Q_i, which cancels with the first term. Thus, the equation reduces to the following, which represents a significant simplification

$$\sum_j n_{i,j} \ln \Gamma_{i,j}^o = Q_i \ln Q_i - \sum_j n_{i,j} q_j \ln \left(\sum_{mj} q_{mj} n_{i,mj} \psi_{mj,j} \right) \tag{D11}$$

Furthermore, we need no longer define $\Theta C_{i,j}$ or $XC_{i,j}$, so that Eqs. C17–C20 are eliminated.

Note

Since we have discarded Eqs. C4 and C6, C13 and C15, C18 and C20 we must introduce the following condition

$$\sum_i X_i^l = 1 \tag{D12}$$

Appendix E: Full Master Problem Objective

The objective in our full master problem is constructed from all constraints in the *full* impact model which include both continuous variables and the integer array n_{ij}. Those which include only continuous variables are redundant since all continuous variables are fixed in the master problem. Thus, the master objective includes only certain UNIFAC equations, namely Eqs. 41, 42, and 44–47.

While Eqs. 41, 42, and 45 are linear in n_{ij}, Eqs. 44, 46, and 47 are not. Thus, these latter equations must be linearized in order to construct the master problem. This is achieved by defining additional terms. Equation 44 illustrates this since sufficient additional terms have already been defined, n_{ij} appears in both $R3_j$ and $R4_{i,j}$ so that Eq. 44 is indeed nonlinear in n_{ij}. However, both $R3_j$ and $R4_{i,j}$ are continuous variables (as, of course, is Q_i), so that in the master these variables are fixed. Thus, Eq. 44 is linearized for the master.

Following the same logic, we introduce two new variables which allow us to treat Eqs. 46 and 47 in the same way. We define

$$S1_j = \sum_{ii} n_{ii,j} X_{ii}^l, \quad \forall j \in J$$

$$S2_{i,j} = \sum_{mj} q_{mj} n_{i,mj} \psi_{mj,j}^{VLE}, \quad \forall i \in I, \forall j \in J$$

so that Eqs. 46 and 47 become

$$R4_{i,j} = \sum_{mj} n_{i,mj} q_{mj} q_j S1_j \psi_{mj,j}^{VLE}, \quad \forall i \in I, \forall j \in J$$

$$R2_i = Q_i \ln Q_i - \sum_j n_{i,j} q_j \ln S2_{i,j}, \quad \forall i \in I$$

Since $S1_j$ and $S2_{i,j}$ are continuous variables they are fixed in the master so that Eqs. 46 and 47 are linearized in the master. Equations 46 and 47 are replaced by the above four equations in the primal so that Lagrange multipliers are generated for each equation. A Lagrange multiplier is needed for each constraint in order to construct the master objective, which is shown in the solution procedure subsection.

Manuscript received July 30, 1998, and revision received Jan. 19, 1999.