Massive, Automated Solvent Screening for Minimum Energy Demand in Hybrid Extraction-Distillation using COSMO-RS

Jan Scheffczyk[a], Christian Redepenning[b], Christian M. Jens[a], Benedikt Winter[a], Kai Leonhard[a], Wolfgang Marquardt[b], André Bardow[a,*]

[a] *RWTH Aachen University, Chair of Technical Thermodynamics, 52056 Aachen, Germany*
[b] *RWTH Aachen University, Process Systems Engineering, 52056 Aachen, Germany*

Abstract

An automated approach for large-scale solvent screening is presented based on a comprehensive process-level assessment. In this solvent screening approach, COSMO-RS is used to efficiently predict physical properties for large numbers of solvents without the need for experimental data. The predicted thermodynamical behavior is used in pinch-based separation models for a thermodynamically sound and robust calculation of the minimum energy demand. With this approach, the performance of a hybrid extraction-distillation process is evaluated fully automated for more than 4,600 solvents. The massive solvent screening approach is successfully applied to purification of the bio-based platform chemical γ-valerolactone (GVL). Novel promising solvents are identified. A reduction of 63 % is achieved in minimum energy demand using the best predicted solvent in comparison to the literature benchmark. Restricting the approach to known classes of solvents, we still find a reduction of 31 %. The process-level assessment overcomes the limitations of heuristics based on physical properties only, and allows for efficient and robust solvent screening.

Keywords: Solvent screening, COSMO-RS, Shortcut process model,

*Corresponding author
Email address: andre.bardow@ltt.rwth-aachen.de (André Bardow)

DOI: 10.1016/j.cherd.2016.09.029

Extraction, Distillation, Rectification Body Method, γ-valerolactone

1. Introduction

The efficiency of chemical processes and thus economics and environmental impacts often crucially depend on the employed solvents [1]. Therefore, solvent selection is an important task in conceptual process design [2]. However, the large number of possible solvents makes solvent selection tedious and computationally demanding. Thus, efficient solvent selection approaches are required. They critically depend on two fundamental features: A sound thermodynamic property prediction and assessment criteria for solvent performance that are reliable, yet computationally tractable.

Current methods for solvent selection commonly employ simplified property models, e.g., first-order group contributions (GC) methods [3] to be computationally tractable. Typically, GC methods require experimental data and thus are restricted to parametrized classes of compounds. To overcome the need for experimentally determined interaction parameters, recent solvent selection approaches focus on the integration of property prediction based on quantum mechanics (QM) [4, 5]. QM-based approaches are, however, computationally demanding in general. Recently, promising approaches of QM-based property prediction have been presented using the efficient thermodynamic model COSMO-RS [6]. Zhou et al. used COSMO-RS for the design of novel components in chemical reactions [7] and integrated solvent/process design [8]. Some of the present authors successfully applied COSMO-RS in an optimization-based computer-aided molecular design (CAMD) framework called COSMO-CAMD [9]. COSMO-RS is generally applicable to systems without experimentally determined GC parameters and thus allows for the efficient evaluation of large numbers of solvents in various mixtures. For these reasons, in this work, COSMO-RS is employed for property prediction in large-scale databank screenings for solvents.

Besides sound thermodynamic property prediction, the quality of the se-

lected solvent depends strongly on the performance assessment criteria used [2]. Currently, solvents are often selected based on single [10] or multi-objective [11, 12] selection criteria. These solvent selection criteria are usually simplified process performance indicators based on physical solvent properties such as selectivity, solvent loss or phase distribution coefficients [13]. The choice of a solvent selection criterion in the pre-selection stage is critical: A suboptimal choice can lead to suboptimal process performance [10].

Papadopoulos and Linke [12] show that several targets for solvent properties exist simultaneously in process flowsheets. In addition, desired solvent properties inherently include trade-offs. E.g., high affinity of the solute to the solvent is desired for efficient extraction, whereas a low affinity between solvent and solute helps to reduce the energy demand in solvent recovery by distillation. Papadopoulos and Linke [12] show by multi-objective optimization (MOO) that these trade-offs in desired solvent properties cannot be captured by evaluating single solvent properties. Preferentially, solvent performance is directly evaluated on the process-level to fully capture the relevant trade-offs in solvent properties [14].

Various approaches for targeting solvent performance on process-level evaluation have been proposed. Bardow et al. [15] identify a hypothetical optimal molecule during process optimization. The resulting so-called continuous-molecular targeting for computer-aided molecular design (CoMT-CAMD) has been successfully applied to identify promising solvent candidates for physical absorption [16] or working fluids in organic Rankine cycles [17]. Similarly, Pereira et al. integrate solvent and process design for methane recovery from carbon dioxide [18]. Other approaches have employed hybrid-stochastic optimization to identify solvent candidates for a coupled absorption-desorption process [19]. A comprehensive review of current CAMD approaches is given by, e.g., Ng. et al. [1]. These studies show that the quality of the solvent selection critically depends on the quality of the process models.

Rigorous process models lead to highly accurate results but are laborious to solve and difficult to automate [20]. In particular, the initialization and thus

the convergence limits the practical application to large sets of solvents. Thus, current conceptual process design approaches reduce the numbers of solvent candidates in pre-selection steps [10]. E.g., Burger et al. [11] extend the approach of Papadopoulos and Linke [12] using simplified process models in the solvent pre-selection stage.

Simplified process models are typically classic shortcut methods such as the well-known equations of Kremser [21] or Underwood [22]. A drawback of these classic shortcut methods models is that they strongly simplify the underlying process model or thermodynamics which can cause inaccuracies, especially for non-ideal separations [23].

In contrast, advanced pinch-based shortcut models provide a thermodynamically sound objective [23, 24]. These pinch-based shortcut models exploit the concept of vanishing thermodynamic driving force in the so-called pinch-points. The identification of controlling pinch-points in each column section significantly simplifies the calculation procedure while still being thermodynamically accurate. The shortcut models assume an infinite number of separation stages and thus operate at a point of minimum reflux or minimum solvent demand [25]. Accordingly, shortcut models give a tight lower bound of the *minimum energy demand* of the process. This class of advanced pinch-based shortcut models was used successfully in various applications, e.g., conceptual process design [24], reaction-separation process design [26], extractive distillation [20], reactive rectification [27], hybrid extraction-distillation [28], multi-component extraction [29] and multi-component absorption [30].

In this work, we use pinch-based shortcut process models for a systematic screening of the performance of solvents in a hybrid extraction-distillation process. Shortcut process models for distillation [31] and extraction [29] are integrated in a process flowsheet. Thereby, a reliable evaluation of the minimum energy demand of the process is possible within seconds. The evaluation of the minimum energy demand is combined with COSMO-RS property prediction in a fully automated solvent screening approach. No pre-selection of solvents based on simplified solvent performance indicators is required but databanks contain-

ing several thousands of solvent candidates can be evaluated within hours by an objective function reflecting the process performance. In the following sections, the proposed approach is presented and challenged in a case study for the hybrid extraction-distillation of the bio-based platform chemical γ-valerolactone (GVL) [32]. In the final section, the results of the solvent screening for minimum energy demand are compared to existing process performance indicators.

2. Methods

2.1. Process Flowsheet

In this work, we consider the purification of a diluted solute A from a carrier C in a hybrid-extraction distillation process [23]. A fixed flowsheet structure is used for the hybrid extraction-distillation. In this flowsheet, the process is limited to a single distillation step and the solute A is recovered as a heavy boiler (Fig. 1).

In the hybrid extraction-distillation process, a feed stream F with the molar composition $\mathbf{z}_\mathrm{F}$ enters an extraction column where a solvent stream S is used to extract the solute A into the extract stream E. The raffinate stream R, deprived of solute A, leaves the extraction column and can further be processed, e.g., by wastewater treatment. The extract stream E is subsequently fed to a distillation column for further purification. In the distillation column, the solute A is separated from the solvent X and co-extracted carrier C. Solvent X and co-extracted carrier C are phase separated in a decanter stage and the solvent rich-phase is recycled to the extraction column. The carrier-rich phase is sent to wasterwater treatment. Solvent loss is replaced by solvent make-up.

The process flowsheet is modeled by advanced pinch-based shortcut process models for extraction [29] and distillation [31]. For the extraction, isothermal separation is assumed and for the distillation isobaric separation. The shortcut process methods assume an infinite number of stages which allows for a sharp split of the components in the separation steps. Thus, solute A is completely extracted into the extract stream E and the raffinate is free of solute A. Sim-

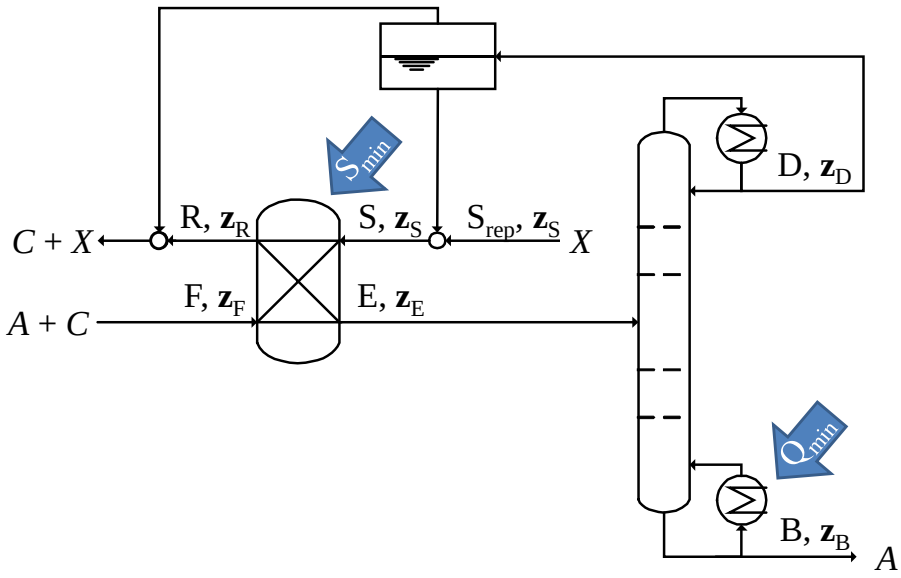

Figure 1: Process flowsheet for hybrid extraction-distillation. Feed stream F containing solute A and carrier C, solvent stream S containing solvent X, raffinate stream R, extract stream E, distillate stream D, bottom stream B, solvent make-up stream $\mathrm{S_{rep}}$ with corresponding compositions $\mathbf{z}_\mathrm{F}$, $\mathbf{z}_\mathrm{S}$, $\mathbf{z}_\mathrm{R}$, $\mathbf{z}_\mathrm{E}$, $\mathbf{z}_\mathrm{D}$, $\mathbf{z}_\mathrm{B}$. $S_{\min}$ is the minimum amount of solvent required for the specified separation task, $Q_{\min}$ is the minimum energy demand required for the specified separation task.

ilarly, the solute A is completely recovered as pure product in the distillation. Consequently, the carrier-rich phase from the phase separation in the decanter stage has an identical composition as the raffinate stream R.

The energy demand in the process is mostly determined by the heat demand of the reboiler in the distillation column. Other heating devices (e.g., preheating of the feed stream to boiling temperature) are assumed to be negligible. Thus, in this work, the process energy demand is represented by the energy requirement for distillation which is expressed by the minimum energy demand $Q_{\min}$. This minimum energy demand $Q_{\min}$ depends on the minimum amount of solvent $S_{\min}$ required for solute extraction in the extraction column. Thus, the performance of a solvent in the process is expressed by a single target function: The minimum energy demand $Q_{\min}$.

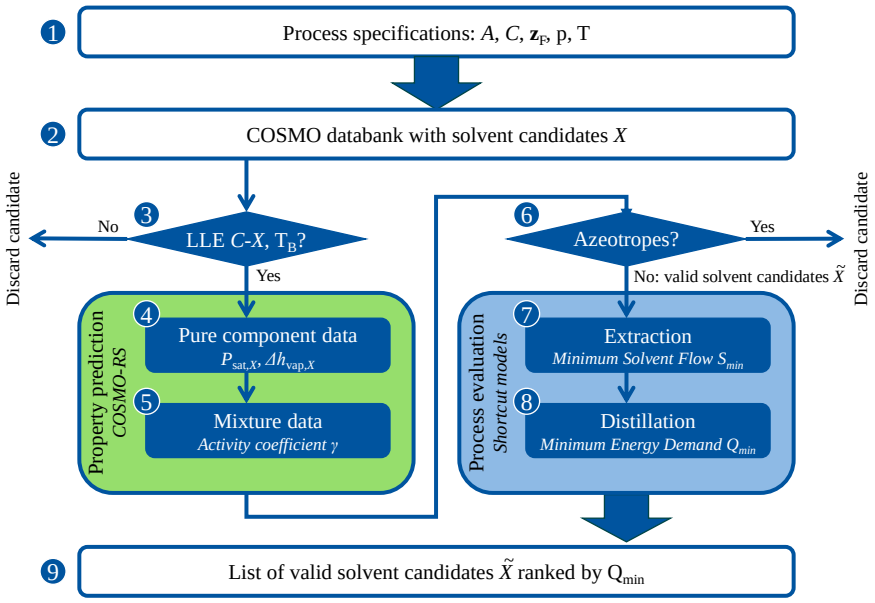

Figure 2: Schematic procedure of the proposed solvent screening approach.

2.2. Solvent Screening

In the proposed solvent screening approach, the solvent X in the hybrid extraction-distillation process in Fig. 1 is considered as a degree of freedom. Thus, the aim of the screening is to identify a solvent candidate X with the lowest minimum energy demand $Q_{\min}$ from a specified set of solvents. The solvent screening consists of the following steps (cf. Fig. 2):

1. Process specifications need to be set. Thus, the solute A and the carrier solution C are specified. In addition, the feed composition $\mathbf{z}_F$ needs to be specified. For the extraction temperature T is specified, for the distillation pressure p is set.
2. Design specifications on the solvent candidates X are imposed. A possible set of solvent candidates is defined by imposing a desired range for molecular size or specifying molecular constituents. E.g., certain groups

of atoms such as halogenes can be excluded from consideration. All solvent candidates meeting these design specifications are selected in the COSMO-RS databank for the screening. The COSMO databank contains so-called COSMO files that result from quantum-mechanic structure optimization. COSMO-RS uses these COSMO files to predict thermodynamic properties. For a detailed description of the COSMO-RS method, see [6].

For each selected solvent candidate X in the COSMO-RS databank, relevant thermodynamic properties are predicted by COSMO-RS:

3. The solvent candidates are tested to meet property constraints that prohibit their application as solvents in a hybrid extraction-distillation process. For this purpose, a preliminary evaluation of specific solvent properties is performed. The existence of a liquid-liquid equilibrium (LLE) in the binary mixture of the solvent X and the carrier C is mandatory for the use of the solvent candidate as an extraction agent. Thus, all solvent candidates are tested for the existence of this LLE using the binary LLE calculation in COSMOtherm [33]. This calculation checks whether an LLE exists. At this point, no constraint is added on the composition of the phases. Solvent candidates without a LLE are discarded. Additionally, solvent candidates are tested to be liquid at ambient temperature and pressure by calculating the boiling point temperature $T_{\mathrm{B},X}$ for all solvent candidates. Two limits are imposed on $T_{\mathrm{B},X}$: First, solvent candidates that are not liquid at ambient temperature with a safety margin of 15 K ($T_{\mathrm{B},X} \leq 313.15\,\mathrm{K}$) are discarded. Second, to obtain pure solute in the bottom product of the distillation column (cf. Sec. 2.1), all solvents that are heavy keys are discarded.
4. For the solvent candidates meeting all requirements from step 3, pure component data is calculated. In particular, Antoine parameters to predict the vapor pressure $p_{\mathrm{sat},X}$ and the molar enthalpy of vaporization $\Delta h_{\mathrm{vap},X}$ are calculated using COSMO-RS (cf. Appendix A). This pure component data is used to estimate boiling temperature and the enthalpy of

vaporization in the process shortcut models.

5. In the next step, mixture data are calculated. Non-idealities in liquid-liquid systems and vapor-liquid systems are described using activity coefficients γ. Isothermal activity coefficients γ can directly be calculated using COSMO-RS. The mixture behaviour of the ternary system *A*-*C*-*X* is described in the shortcut process models using parameters of the *non-random-two-liquid* (NRTL) [34] model. NRTL parameters of the binary systems *A*-*C*, *A*-*X* and *C*-*X* are used to predict phase equilibria in the ternary system *A*-*C*-*X* over a specified range of temperature. Thus, a regression for NRTL parameters based on isothermal activity coefficients γ is performed. (cf. Appendix A).
6. To limit the purification by distillation to a single distillation column, no ternary azeotropes and azeotropes between solvent and solute can be present. Thus, all azeotropes in the ternary system are calculated based on the algorithm proposed by Fidkowski et al. [35]. Solvent candidates with ternary azeotropes and azeotropes between solvent and solute are discarded.

All solvent candidates X that meet the requirements from step 1 to step 6 are considered valid solvent candidates $\widetilde{\mathrm{X}}$. For all valid solvent candidates, the process model is evaluated:

7. For each valid solvent candidate, the extraction unit operation is evaluated. The extraction unit operation is modeled using the shortcut process model and solved using the procedure proposed by Redepenning et al. [29]. For a given feed composition, the procedure calculates the minimum amount of solvent S_{min}, as well as all outlet compositions and flow rates (cf. Fig. 1).
8. Next, for each valid solvent candidate $\widetilde{X}$, the distillation unit operation is evaluated. The distillation unit operation is modeled using the rectification body method (RBM) proposed by Bausa et al. [31]. Limitations by continuous molecular overflow assumption are overcome by considering

the enthalpy of vaporization in the RBM shortcut process model. Further contributions to the energy demand due to temperature changes in vapor and liquid phases are assumed to be negligible. Thus, no molar heat capacities are required in the calculation. The RBM shortcut process model returns the minimum energy demand $Q_{\min}$. All process streams for the hybrid extraction-distillation process with all corresponding compositions (cf. Fig. 1) are now determined.

9. As a result, a list is returned containing all valid solvent candidates $\widetilde{X}$ ranked by their respective minimum energy demand $Q_{\min}$.

3. Case study: Purification of the platform chemical γ-valerolactone

In this section, the proposed screening approach (Sec. 2.2) is exemplified for the identification of novel solvents for the purification of γ-valerolacton (GVL). GVL is an intermediate in the production of bio-based value products and is a promising precursor for fuel and commodity chemicals [32]. GVL can be produced from lignocellulosic biomass and needs to be recovered and purified from aqueous solutions for further processing. A recent approach for the recovery of GVL is a hybrid extraction-distillation process proposed by Murat Sen et al. [36]. In their work, Murat Sen et al. use n-butyl acetate as extraction solvent which is considered as benchmark is this case study.

A massive databank screening of more than 4,600 solvent candidates is performed to identify novel promising solvents for GVL purification. The minimum energy demand $Q_{\min}$ is used as the target function.

For the solvent screening, the following process specifications are set (step 1). A feed composition of $\mathbf{z}_{\mathrm{F}} = (0.05, 0.95, 0.00)$ for the ternary system GVL-water-X is assumed. Pressure and temperature in the process are set to $p = 1\,\mathrm{bar}$ and $T = 25\,^{\circ}\mathrm{C}$. Bio-compatibility of the solvent candidates X is desirable, thus, possible solvents candidates X are limited to molecules only containing carbon, hydrogen and oxygen atoms (step 2).

All considered solvent candidates are tested to meet the required property

constraints (step 2). For 4,331 solvent candidates X in the databank, a binary LLE in the system water-X exists. Additionally, the boiling temperature $T_{\mathrm{B},X}$ is evaluated. Solvents with a boiling temperature higher than the predicted boiling temperature of GVL ($T_{\mathrm{B,GVL}} = 507\,\mathrm{K}$) are discarded. Overall, 1,715 solvent candidates also meet the required property constraints on the boiling temperature $T_{\mathrm{B},X}$ and are further evaluated. For those solvent candidates, pure- and mixture-data are computed (step 4 and 5). The computational time to evaluate all 1,715 solvent candidates X is $\approx 14\,\mathrm{h}$ on a 3.2 GHz desktop computer using 4 parallel cores. This computational time demand is almost exclusively due to the computation demand of COSMO-RS property prediction. In particular, the generation of (temperature-dependent) NRTL parameters is computationally expensive. The calculations can easily be accelerated using parallel computing. In contrast to COSMO-RS property prediction, the process shortcut models are evaluated within minutes ($\approx 5\,\mathrm{min}$). The calculated NRTL parameters are stored and can be re-used. E.g., for the ternary system A-C-X, three sets of binary NRTL parameters are calculated: A-C, A-X and C-X. Binary NRTL parameters for the system A-C are independent from the solvent candidates X and have to be calculated only once. In addition, binary NRTL parameters for the systems C-X are independent from the solute A. Since the binary NRTL parameters for the systems C-X have already been calculated in this work for C = water, screening for a different solute A in water would require $\approx 50\,\%$ less time: Only binary NRTL parameters for the system A-X would have to be re-calculated for all solvent candidates X.

For all 1,715 solvent candidates X, azeotropes are calculated and evaluated (step 6). Overall, 1,439 solvent candidates X meet the requirements from (step 1 to 6) and thus are regarded valid solvent candidates $\widetilde{X}$. For all valid solvent candidates $\widetilde{X}$, the hybrid extraction-distillation process is evaluated (step 7 and 8). As a result, all 1,439 valid solvent candidates are ranked by the corresponding minimum energy demand Q_{min}.

Fig. (3) shows the result of the solvent screening. Promising solvent candidates are identified in Tab. 1. In particular, 155 solvents are predicted to have

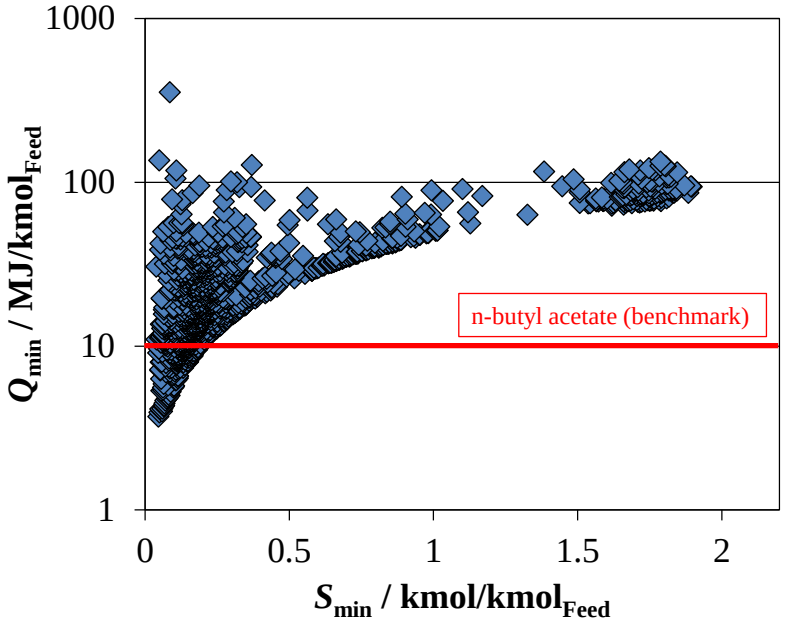

Figure 3: Minimum energy demand $Q_{\min}$ and minimum amount of solvent $S_{\min}$ for 1,439 valid solvent candidates $\widetilde{X}$ (blue diamonds) and $Q_{\min}$ for benchmark solvent n-butyl acetate (red line).

a lower minimum energy demand $Q_{\min}$ than the literature benchmark n-butyl acetate. Tab. 1 shows that the top 5 identified solvents are aliphatic alkynes. The molecule with the lowest minimum energy demand is 1,5-hexadiyne with 63 % reduction of $Q_{\min}$ (67 % reduction of $S_{\min}$) in comparison to the benchmark solvent n-butyl acetate. Thus, 1,5-hexadiyne is highly promising in terms of the energy saving potential.

All top solvents identified in Tab. 1 have lower $Q_{\min}$ and lower $S_{\min}$ compared to the benchmark. This indicates that the extraction contributes significantly to the good performance of the top solvents in the distillation. A comparison of the ternary diagram of n-butyl acetate (benchmark) and 1,5-hexadiyne (best solvent identified) reveals the excellent extraction agent properties of 1,5-hexadiyne. Fig. (4) shows a larger miscibility gap and steeper tie lines for 1,5-hexadiyne

Table 1: Selected results for solvent screening using shortcut model and reduction of minimum energy demand compared to benchmark (BM). Available experimental boiling point temperature $T_{\mathrm{B},X}$ (exp.) and melting point $T_{\mathrm{Melt},X}$ (exp.) are provided for further solvent assessment.

Solvent	S_{min}	Q_{min}	Q_{min} reduction	Rank	$T_{\mathrm{Melt},X}$ (exp.)	Ref.	$T_{\mathrm{B},X}$ (exp.)	Ref.
	$\mathrm{kmol/kmol_{Feed}}$	$\mathrm{MJ/kmol_{Feed}}$	%		°C		°C	
1,5-hexadiyne	0.05	3.70	63	1	-4	[37]	88	[37]
1,6-heptadiyne	0.05	3.94	61	2	-85	[37]	112	[37]
1,3-hexadien-5-yne	0.06	3.99	61	3			82	[38]
1,7-octadiyne	0.05	4.12	59	4			136	[39]
1-penten-4-yne	0.06	4.14	59	5			42	[40]
furan	0.07	4.64	54	9	-86	[41]	31	[41]
3-methylfuran	0.09	5.26	48	15			65	[42]
2-methylfuran	0.11	6.36	37	27			64	[37]
2,3-hexanedione	0.10	6.98	31	34			128	[43]
3,4-hexanedione	0.11	7.15	29	36				
n-butylacetate	0.14	10.11	-	156	-74	[37]	126	[37]

compared to the benchmark. Thus, a high selectivity and a high capacity for the solute GVL are achieved which results in a low minimum amount of solvent S_{min}. In turn, a low S_{min} consequentially leads to a reduced extract stream E which reduces the minimum energy demand in the distillation Q_{min}.

Rigorous process model calculations using ASPEN Plus V8.4 are performed to validate the results from shortcut process models. The minimum energy demand in ASPEN Plus $Q_{\mathrm{min,ASPEN}}$ is calculated for a flowsheet according to Fig. (1). The flowsheet is initialized with results from the shortcut process models and converged to minimum solvent flow and minimum reboiler heat duty (for calculation details see SI).

Overall, the results of the shortcut process models are in good agreement with the rigorous process calculations. The mean average percentage error for the top 50 solvents is +9 %. In particular, the minimum energy demand in ASPEN Plus for the best solvent identified 1,5-hexadiyne is $Q_{\mathrm{min,ASPEN}} = 3.97\,\mathrm{MJ/kmol_{Feed}}$ (+7 %) and for the benchmark n-butyl acetate $Q_{\mathrm{min,ASPEN}} = 11.10\,\mathrm{MJ/kmol_{Feed}}$ (+9 %). Additionally, the high selectivity and capacity of 1,5-hexadiyne for GVL is confirmed by the low minimum solvent demand in ASPEN Plus ($S_{\mathrm{min,ASPEN}} =$

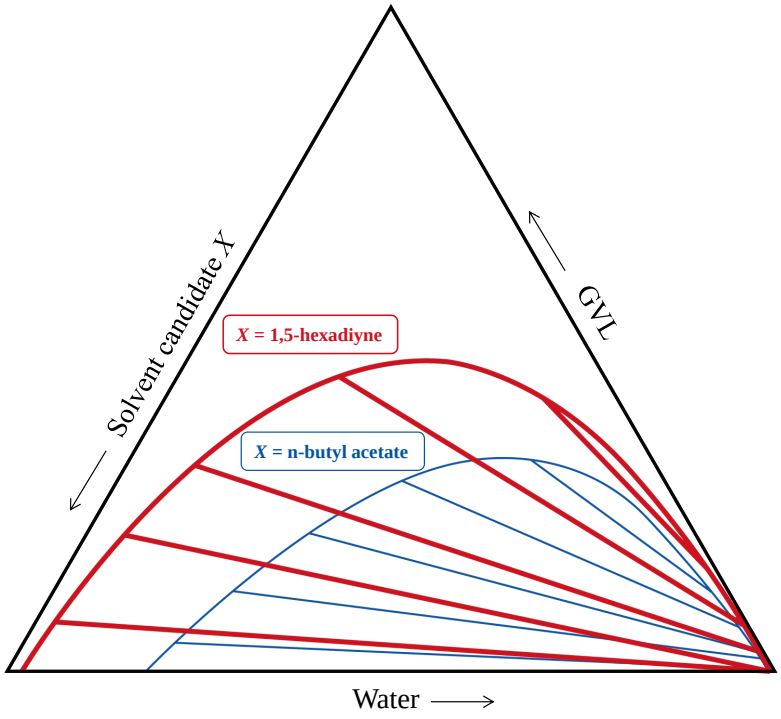

Figure 4: Ternary diagram of literature benchmark n-butyl acetate (blue thin line) and best solvent identified by solvent screening 1,5-hexadiyne (red thick line).

0.05 kmol/kmol$_{\text{Feed}}$). Thus, rigorous process model calculations in ASPEN Plus validate the predicted potential to reduce the minimum energy demand and determine a reduction by 64 % with 1,5-hexadiyne. This validation supports the presented approach for identification of promising solvents by the shortcut process models. An extended list with validated results for the top 50 solvents is provided in the SI. The solvent ranking by minimum energy demand generated by the screening approach and the ASPEN Plus simulations are in very good agreement with a Spearman's rank correlation coefficient of 0.93, where an ideal correlation corresponds to a rank correlation coefficient of 1 [44]. However, the ASPEN Plus simulations also show that two components, trans-3-penten-1-yne and furan, ranked #7 and #9 by the screening (Tab. 1) require even less energy demand than 1,5-hexadiyne according to the ASPEN Plus simulations. This effect is due to the different settings in the flowsheets (lower convergence tolerance and finite number of stages in ASPEN Plus, see SI). This finding also highlights the importance to not focus only on the top candidate from the screening but to consider the list of top candidates generated by our approach.

Only 59 of the 100 best solvents could have been evaluated using the modified UNIFAC (Dortmund) group-contribution method [45]. This is mainly caused by the lack of parametrized groups for promising furanes and missing interaction parameters for alkynes with water. This highlights a key feature of the presented approach that is independent of experimentally parametrized group interaction parameters.

To be a solvent of practical relevance, further criteria need to be evaluated next to minimum energy demand, e.g., design limits on melting point temperature or boiling point temperature as well as toxicity and chemical stability. These criteria are currently evaluated by human post-processing. For the present case study, we find that all available experimental melting point temperatures indicate that the selected solvents are liquids at room temperature (Tab. 1). Similarly, experimental boiling points in Tab. 1 show that all components are within the desired design limits on boiling point temperature. We further find that alkynes have a tendency to decompose and are usually stored under low

temperatures to prevent degradation. Thus, the alkynes ranked as top solvents candidates in Tab. 1 seem questionable for practical application. However, an advantage of the proposed solvent screening approach is that a ranked list of solvents is generated (extended list of all top 155 solvents can be found in the SI). Further criteria for practical relevance can be evaluated by the design engineer.

For practical relevance, promising solvents are diketones and furanes which are already discussed in the literature as fuel candidates and solvents [46]. Tab. 1 shows that promising furanes (#9, #15, #27) are identified in the solvent screening, which are stable under relevant process conditions according to manufacturer data. The identified furanes reduce the minimum energy demand Q_{min} by $\approx$ 40-50 %. The experimental boiling point temperature of furan is $T_{\mathrm{B,furan}} = 31\,^{\circ}\mathrm{C}$ (Tab. 1) which is close to ambient temperature. Here, solvent loss by evaporation should be considered for practical applications. Future screening studies could add this constraint explicitly to the design problem. If further criteria such as toxicity are applied, 2,3-hexanedione (#34) seems to be a very promising candidate for GVL extraction. In contrast to the furanes, this diketone is non-toxic according to manufacturer data. 2,3-hexanedione reduces the minimum energy demand Q_{min} by 31 % in comparison to the benchmark and is thus proposed as the most promising solvent from a practical perspective. Additionally, 2,3-hexanedione is commercially available and can further be evaluated experimentally.

4. Comparison to conventional process performance indicators

In this section, the results from solvent screening based on minimum energy demand are compared to conventional process performance indicators. All identified 1,439 valid solvent candidates $\widetilde{X}$ are considered for the comparison. The following conventional screening criteria are evaluated (for calculation details see Appendix B):

1. Phase distribution coefficient P (at infinite dilution of solute A) is commonly used to assess the solvent extraction selectivity (e.g., [47])

2. Relative volatility α is commonly used to assess the energy demand in the distillation (e.g., [13])

To quantify the correlation between conventional process indicators and the results from the shortcut process models, the Pearson correlation coefficient $r_{a,b}$ is evaluated. The Pearson correlation coefficient determines the correlation of sample values a_i and b_i. The Pearson correlation coefficient $r_{a,b}$ can take on values from $r_{a,b} = -1$ to $r_{a,b} = 1$. Pearson correlation coefficients of $r_{a,b} = -1$ and $r_{a,b} = 1$ correspond to an ideal linear correlation of a_i and b_i. In contrast, $r_{a,b} = 0$ indicates no correlation.

Fig. (5) shows a comparison of conventional process performance indicators to the results obtained from the screening approach based on minimum energy demand.

A very good correlation is found between phase distribution coefficient P and S_min (Fig. 5A). This is reflected in the high Pearson correlation coefficient $r_{\log P, \log S_\mathrm{min}} = -0.97$. Accordingly, all top solvents (green triangles in Fig. (5A)) achieve high P values. Similarly, Fig. 5B shows that P is moderately good correlated to the minimum energy demand Q_min ($r_{\log P, \log Q_\mathrm{min}} = -0.76$). A solvent screening based on phase distribution coefficient P would identify 2-methylphenol as the best solvent. 2-methylphenol can be considered a good solvent candidate since it has a lower minimum energy demand Q_min than the benchmark molecule. However, a solvent assessment based on the phase distribution coefficient P has significant drawbacks. Firstly, the best solvent identified based on the phase distribution coefficient P has a Q_min of $9.63\,\mathrm{MJ/kmol_{Feed}}$ which is $\approx 160\,\%$ higher than the best solvent found in the screening based on shortcut process models. The reason for the high minimum energy demand is the comparably high boiling point temperature predicted for 2-methylphenol $T_\mathrm{B,2\text{-}methylphenol} = 201\,^\circ\mathrm{C}$, which is in agreement with experimental findings (experimental $T_\mathrm{B,2\text{-}methylphenol} = 192\,^\circ\mathrm{C}$ [48]). Secondly, Fig. (5B) shows that a screening based on the phase distribution coefficient P does not identify *all* solvents with a lower minimum energy demand Q_min than the benchmark. Thus,

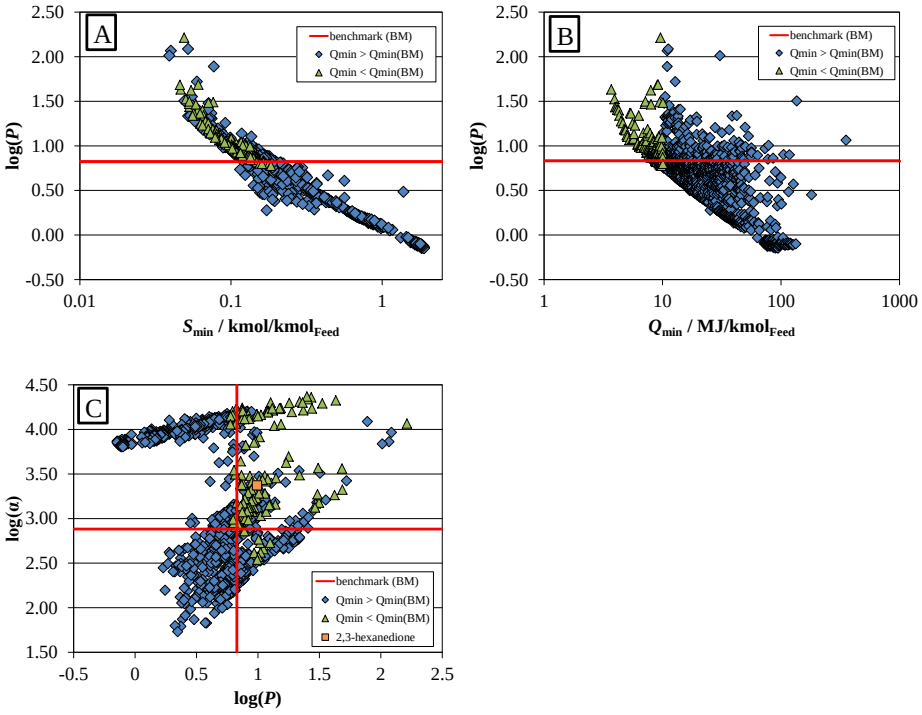

Figure 5: Comparison of phase distribution coefficient P to to minimum amount of solvent $S_{\min}$ (A) and minimum energy demand $Q_{\min}$ (B). Solvent screening based on phase distribution coefficient P and relative volatility α (C). Red lines: benchmark (BM), green triangles: promising solvents ($Q_{\min} < Q_{\min}$(BM)), blue diamonds: other solvents ($Q_{\min} > Q_{\min}$ (BM)), orange square: most promising identified solvent 2,3-hexanedione.

promising solvents with a low phase distribution coefficient P would be missed in a solvent selection based on the phase distribution coefficient P.

Further refinement of solvent selection based on conventional process indicators commonly employs the relative volatility α ([13]). The best identified solvent based on α is 2-methyl-1-buten-3-yne with a $Q_{\mathrm{min}} = 4.28\,\mathrm{MJ/kmol_{Feed}}$, which is 16 % higher than the best solvent found in the screening based on shortcut process models. The good performance of 2-methyl-1-buten-3-yne can be explained by its low boiling point ($T_{\mathrm{B,2\text{-}methyl\text{-}1\text{-}buten\text{-}3\text{-}yne}} = 51\,^{\circ}\mathrm{C}$) and low enthalpy of vaporization ($\Delta h_{\mathrm{vap,2\text{-}methyl\text{-}1\text{-}buten\text{-}3\text{-}yne}} =$ of $31.8\,\mathrm{kJ\,mol^{-1}}$). The low correlation of the relative volatility α and the minimum energy demand Q_{min} (correlation coefficient $r_{\log\alpha,\log Q_{\mathrm{min}}} = 0.24$) does not allow for an evaluation of the solvent process performance solely based on α.Yet, relative volatility α seems to be a good heuristic process performance indicator for further solvent selection in combination with distribution coefficient P (Fig. 5C). In a solvent screening based on both conventional process indicators P and α, most solvents with lower minimum energy demand Q_{min} than the benchmark are properly identified as promising. However, a solvent selection based on process indicators P and α still comes with major drawbacks: First, not all solvents with lower minimum energy demand Q_{min} than the benchmark are identified: Overall, 16 % of all promising solvent candidates are missed in a solvent selection based on P and α. Moreover, 116 components with a higher minimum energy demand Q_{min} than the benchmark are falsely identified as promising. Further conventional process performance indicators, i.e., constraints on the boiling point temperature, would thus be necessary to rule out more candidates which comes at the risk of losing further promising solvent candidates. Second, there is no direct correlation between the conventional process indicators and Q_{min}, i.e., it is not known whether favorable α is more important than favorable P. Importantly, many promising solvents with low Q_{min} are not located at extreme values of P or α but at intermediate values of P or α (cf. Fig. 5C). In particular, the solvents of practical relevance discussed above, e.g., 2,3-hexanedione with $\log P = 0.99$ and $\log\alpha = 3.37$, are identified at intermediate values of P and α (orange square

in Fig. (5C)). Thus, for a reliable identification of all promising solvents, a quantitative ranking of the solvents is necessary. This quantitative ranking by minimum energy demand $Q_{\min}$ is provided by the shortcut process models.

5. Conclusions

A massive solvent screening approach based on COSMO-RS and pinch-based shortcut process models is presented. Automated evaluation of pinch-based process models for extraction and rectification devices was established to screen solvents by the minimum energy demand. The approach is not limited to a reduced number of components which is highlighted for a fully automated solvent screening of a large-scale databank.

Results show that more than 4,600 solvents can be screened fast and efficiently with the presented screening approach. Novel promising solvents for the hybrid extraction-distillation of GVL are identified with predicted better performance than the literature benchmark n-butyl acetate. A theoretical reduction of the minimum energy demand $Q_{\min}$ of 63 % is predicted for 1,5-hexadiyne and 31 % for 2,3-hexanedione. The most promising solvent candidate for practical application, 2,3-hexanedione, is commercially available and can further be evaluated experimentally.

A comparison to conventional screening criteria provides insight in the inherent trade-off of desired solvent properties in the process whereas heuristic criteria based on phase distribution coefficient or relative volatility lead to suboptimal solutions. The proposed screening approach captures this trade-off and yields quantitative information on process performance. Thus, the presented screening approach provides a comprehensive process-level assessment of the screened solvents and successfully takes into account inherent process trade-offs.

Overall, the proposed solvent screening approach efficiently combines COSMO-RS property prediction with a comprehensive process-level assessment. Using COSMO-RS, quantum mechanics-based property prediction is achieved overcoming the need for experimental data. The shortcut process models allow for

a fast and reliable process-level evaluation. Thus, the proposed screening approach significantly enlarges the range of current solvent selection approaches and increases confidence in solvent pre-selection.

Acknowledgments

This work was performed as part of the Cluster of Excellence 'Tailor-Made Fuels from Biomass', which is funded by the Excellence Initiative by the German federal and state governments to promote science and research at German universities.

References

References

[1] L. Y. Ng, F. K. Chong, N. G. Chemmangattuvalappil, Challenges and opportunities in computer-aided molecular design, Computers & Chemical Engineering 81 (2015) 115–129.

[2] C. S. Adjiman, A. Galindo, G. Jackson, Molecules matter: The expanding envelope of process design, in: M. R. Eden (Ed.), Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design, Vol. 34 of Comput. aided Chem. Eng., Elsevier, Amsterdam, 2014, pp. 55–64.

[3] J. Gmehling, Present status and potential of group contribution methods for process development, The Journal of Chemical Thermodynamics 41 (6) (2009) 731–747.

[4] H. Struebing, Z. Ganase, P. G. Karamertzanis, E. Siougkrou, P. Haycock, P. M. Piccione, A. Armstrong, A. Galindo, C. S. Adjiman, Computer-aided molecular design of solvents for accelerated reaction kinetics, Nature chemistry 5 (11) (2013) 952–957.

[5] A. Lehmann, C. D. Maranas, Molecular design using quantum chemical calculations for property estimation, Industrial & Engineering Chemistry Research 43 (13) (2004) 3419–3432.

[6] A. Klamt, F. Eckert, W. Arlt, COSMO-RS: an alternative to simulation for calculating thermodynamic properties of liquid mixtures, Annual Review of Chemical and Biomolecular Engineering 1 (2010) 101–122.

[7] T. Zhou, K. McBride, X. Zhang, Z. Qi, K. Sundmacher, Integrated solvent and process design exemplified for a Diels-Alder reaction, AIChE Journal 61 (1) (2015) 147–158.

[8] T. Zhou, Z. Lyu, Z. Qi, K. Sundmacher, Robust design of optimal solvents for chemical reactions—a combined experimental and computational strategy, Chemical Engineering Science 137 (2015) 613–625.

[9] J. Scheffczyk, L. Fleitmann, A. Schwarz, M. Lampe, A. Bardow, K. Leonhard, COSMO-CAMD: A framework for optimization-based computer-aided molecular design using COSMO-RS, Chemical Engineering Science, accepted (2016).

[10] S. Kossack, K. Kraemer, R. Gani, W. Marquardt, A systematic synthesis framework for extractive distillation processes, Chemical Engineering Research and Design 86 (7) (2008) 781–792.

[11] J. Burger, V. Papaioannou, S. Gopinath, G. Jackson, A. Galindo, C. S. Adjiman, A hierarchical method to integrated solvent and process design of physical CO2 absorption using the SAFT-γ Mie approach, AIChE Journal 61 (10) (2015) 3249–3269.

[12] A. I. Papadopoulos, P. Linke, Multiobjective molecular design for integrated process-solvent systems synthesis, AIChE Journal 52 (3) (2006) 1057–1070.

[13] E. J. Pretel, P. A. López, S. B. Bottini, E. A. Brignole, Computer-aided molecular design of solvents for separation processes, AIChE Journal 40 (8) (1994) 1349–1360.

[14] A. I. Papadopoulos, P. Linke, A unified framework for integrated process and molecular design, Chemical Engineering Research and Design 83 (6) (2005) 674–678.

[15] A. Bardow, K. Steur, J. Gross, Continuous-molecular targeting for integrated solvent and process design, Industrial & Engineering Chemistry Research 49 (6) (2010) 2834–2840.

[16] M. Stavrou, M. Lampe, A. Bardow, J. Gross, Continuous molecular targeting–computer-aided molecular design (CoMT-CAMD) for simultaneous process and solvent design for CO2 capture, Industrial & Engineering Chemistry Research 53 (46) (2014) 18029–18041.

[17] M. Lampe, M. Stavrou, J. Schilling, E. Sauer, J. Gross, A. Bardow, Computer-aided molecular design in the continuous-molecular targeting framework using group-contribution PC-SAFT, Computers & Chemical Engineering 81 (2015) 278–287.

[18] F. Pereira, E. Keskes, A. Galindo, G. Jackson, C. Adjiman, Integrated solvent and process design using a SAFT-VR thermodynamic description: High-pressure separation of carbon dioxide and methane, Computers & Chemical Engineering 35 (3) (2011) 474–491.

[19] T. Zhou, Y. Zhou, K. Sundmacher, A hybrid stochastic–deterministic optimization approach for integrated solvent and process design, Chemical Engineering Science (2016) (2016).

[20] S. Brüggemann, W. Marquardt, Shortcut methods for nonideal multicomponent distillation: 3. extractive distillation columns, AIChE Journal 50 (6) (2004) 1129–1149.

[21] B. D. Smith, W. K. Brinkley, General short-cut equation for equilibrium stage processes, AIChE Journal 6 (3) (1960) 446–450.

[22] Underwood, A. J. V., Fractional distillation of multicomponent mixtures, Industrial & Engineering Chemistry 41 (12) (1949) 2844–2847.

[23] M. Skiborowski, A. Harwardt, W. Marquardt, Conceptual design of distillation-based hybrid separation processes, Annual review of chemical and biomolecular engineering 4 (2013) 45–68.

[24] W. Marquardt, S. Kossack, K. Kraemer, A framework for the systematic design of hybrid separation processes, Chinese Journal of Chemical Engineering 16 (3) (2008) 333–342.

[25] S. Brüggemann, W. Marquardt, Conceptual design of distillation processes for mixtures with distillation boundaries: I. computational assessment of split feasibility, AIChE Journal 57 (6) (2011) 1526–1539.

[26] S. Recker, M. Skiborowski, C. Redepenning, W. Marquardt, A unifying framework for optimization-based design of integrated reaction–separation processes, Computers & Chemical Engineering 81 (2015) 260–271.

[27] J. W. Lee, S. Brüggemann, W. Marquardt, Shortcut method for kinetically controlled reactive distillation systems, AIChE Journal 49 (6) (2003) 1471–1487.

[28] K. Kraemer, A. Harwardt, R. Bronneberg, W. Marquardt, Separation of butanol from acetone–butanol–ethanol fermentation by a hybrid extraction–distillation process, Computers & Chemical Engineering 35 (5) (2011) 949–963.

[29] C. Redepenning, S. Recker, W. Marquardt, Pinch-based shortcut method for the design of isothermal extraction columns, in preparation (2016).

[30] C. Redepenning, W. Marquardt, Pinch-based shortcut method for the conceptual design of isothermal absorption columns, in preparation (2016).

[31] J. Bausa, R. v. Watzdorf, W. Marquardt, Shortcut methods for nonideal multicomponent distillation: I. simple columns, AIChE Journal 44 (10) (1998) 2181–2198.

[32] D. M. Alonso, S. G. Wettstein, J. A. Dumesic, Gamma-valerolactone, a sustainable platform molecule derived from lignocellulosic biomass, Green Chemistry 15 (3) (2013) 584.

[33] COSMOlogic GmbH & Co KG, COSMOtherm, C3.0, release 1501. URL http://www.cosmologic.de

[34] H. Renon, J. M. Prausnitz, Local compositions in thermodynamic excess functions for liquid mixtures, AIChE Journal 14 (1) (1968) 135–144.

[35] Z. T. Fidkowski, M. F. Malone, M. F. Doherty, Computing azeotropes in multicomponent mixtures, Computers & Chemical Engineering 17 (12) (1993) 1141–1155.

[36] S. Murat Sen, C. A. Henao, D. J. Braden, J. A. Dumesic, C. T. Maravelias, Catalytic conversion of lignocellulosic biomass to fuels: Process development and technoeconomic evaluation, Chemical Engineering Science 67 (1) (2012) 57–67.

[37] Z. Rappoport, CRC handbook of tables for organic compound identification, 3rd Edition, CRC Press, Boca Raton, Florida, 1967.

[38] F. Sondheimer, D. A. Ben-Efraim, Y. Gaoni, Unsaturated macrocyclic compounds. XVIII. the prototropic rearrangement of linear 1,5-diynes to conjugated polyen-ynes, Journal of the American Chemical Society 83 (7) (1961) 1682–1685.

[39] J. L. Everett, G. A. R. Kon, 610. the preparation of some cytotoxic epoxides, Journal of the Chemical Society (Resumed) (1950) 3131.

[40] Y. N. Bubnov, A. V. Tsyban', B. M. Mikhailov, A convenient synthesis of 1,4-pentenynes (allylacetylenes) via allylboranes, Synthesis 11 (1980) 904–905.

[41] L. Gontrani, F. Ramondo, R. Caminiti, Furan and thiophene in liquid phase: An X-ray and molecular dynamics study, Chemical Physics Letters 422 (1-3) (2006) 256–261.

[42] M. Okabe, H. Tamagawa, M. Tada, Synthesis of (3-furanyl) methyl derivative, Synthetic Communications 13 (5) (2006) 373–378.

[43] I. M. Heilbron, E. R. H. Jones, P. Smith, B. C. L. Weedon, 16. researches on acetylenic compounds. Part IV. The hydration of some acetylenylcarbinols derived from $\alpha\beta$-unsaturated aldehydes, Journal of the Chemical Society (1946) 54–58.

[44] C. Spearman, The proof and measurement of association between two things. By C. Spearman, 1904, The American journal of psychology 100 (3-4) (1987) 441–471.

[45] A. Jakob, H. Grensemann, J. Lohmann, J. Gmehling, Further development of modified UNIFAC (Dortmund): Revision and Extension 5, Industrial & Engineering Chemistry Research 45 (23) (2006) 7924–7933.

[46] I. Delidovich, K. Leonhard, R. Palkovits, Cellulose and hemicellulose valorisation: an integrated challenge of catalysis and reaction engineering, Energy & Environmental Science 7 (9) (2014) 2803–2830.

[47] M. Hostrup, P. M. Harper, R. Gani, Design of environmentally benign processes: integration of solvent design and separation process synthesis, Computers & Chemical Engineering 23 (10) (1999) 1395–1414.

[48] J. M. Brittain, de la Mare, Peter B. D., P. A. Newman, Electrophilic substitution with rearrangement. Part 9. Dienones derived from brominations of o-, m-, and p-cresol, Journal of the Chemical Society, Perkin Transactions 2, 1 (1981) 32.

Appendix A. Property prediction using COSMO-RS

All thermodynamic properties in this work are predicted by COSMO-RS [6]. COSMO files are used from the 1501-BP-TZVP databank and COSMO-RS calculations are performed using COSMOthermX Version 15.

Pure component properties used in this work are vapor pressure, $p_{\mathrm{sat},X}$, and molar enthalpy of vaporization, $\Delta h_{\mathrm{vap},X}$, for each solvent candidate X. The vapor pressure is calculated using the Antoine equation

$$p_{\mathrm{sat},X} = \exp\left(A_X - \frac{B_X}{C_X + T}\right). \tag{A.1}$$

The solvent specific coefficients, A_X, B_X and C_X, are calculated by COSMO-RS.

The molar enthalpy of vaporization, $\Delta h_{\mathrm{vap},X}$, is estimated by combining the Antoine equation, Eq. (A.1), with Clausius-Clapeyron

$$\frac{d \ln\left(p_{\mathrm{sat},X}/p^0_{\mathrm{sat},X}\right)}{dT} = \frac{\Delta h_{\mathrm{vap},X}}{RT^2}, \tag{A.2}$$

which leads to

$$\Delta h_{\mathrm{vap},X} = \frac{B_X}{(T + C_X)^2} RT^2. \tag{A.3}$$

Non-ideal liquid mixture properties are expressed using activity coefficients $\gamma_i(\mathbf{z}, T)$ that are calculated by COSMO-RS. With $\gamma_i(\mathbf{z}, T)$, vapor-liquid equilibria (VLE) and liquid-liquid equilibria (LLE) are calculated.

In this work, activity coefficients $\gamma_i(\mathbf{z}, T)$ are expressed by the non-random-two-liquid NRTL model [34].

The extraction shortcut process model uses isothermal binary NRTL parameters to calculate $\gamma_i(\mathbf{z})$. Isothermal binary NRTL parameters τ_{ij}, τ_{ji}, α_{ij} and α_{ji} are defined by

$$\ln \gamma_1(\mathbf{z}) = z_2^2 \left[\tau_{21} \left(\frac{G_{21}}{z_1 + z_2 G_{21}} \right)^2 + \frac{\tau_{12} G_{12}}{(z_2 + z_1 G_{12})^2} \right], \tag{A.4}$$

$$\ln \gamma_2(\mathbf{z}) = z_1^2 \left[\tau_{12} \left(\frac{G_{12}}{z_2 + z_1 G_{12}} \right)^2 + \frac{\tau_{21} G_{21}}{(z_1 + z_2 G_{21})^2} \right], \tag{A.5}$$

$$\ln G_{12} = -\alpha_{12} \tau_{12}, \tag{A.6}$$

$$\ln G_{21} = -\alpha_{21} \tau_{21}. \tag{A.7}$$

Generally, $\alpha_{12} = \alpha_{21} = \alpha$ is assumed.

In this work, we use an equally-spaced concentration grid for $\mathbf{z} \in [0, 1]$ with $n_{\mathbf{z}} = 11$ evaluations of $\gamma(\mathbf{z})$ to determine τ_{ij}, τ_{ji}, α_{ij} and α_{ji}.

The distillation shortcut process model uses temperature-dependent binary NRTL parameters to calculate $\gamma_i(\mathbf{z}, T)$. In this work, temperature-dependent binary NRTL parameters are calculated using polynomial coefficients a_{ij}, b_{ij}, c_{ij}, d_{ij}, e_{ij} which are defined by

$$\tau_{ij} = a_{ij} + \frac{T}{b_{ij}} + c_{ij} \ \ln T + d_{ij} T, \tag{A.8}$$

$$\alpha_{ij} = e_{ij} + f_{ij} T. \tag{A.9}$$

To determine the parameters a_{ij}, b_{ij}, c_{ij}, d_{ij}, e_{ij}, in this work, we use an equally-spaced temperature grid for $T \in [298.15\,\mathrm{K}, 473.15\,\mathrm{K}]$ with $n_T = 10$ temperature grid points. Overall, $n_{\mathbf{z}} = 11$ evaluations $\gamma(\mathbf{z})$ on each of the $n_T = 10$ temperature levels are performed, i.e., 110 evaluations for each solvent.

Appendix B. Conventional screening criteria

The phase distribution coefficient P is defined as

$$P = \frac{\mathbf{z}_{\mathrm{GVL,E}}}{\mathbf{z}_{\mathrm{GVL,R}}} = \frac{\gamma(\mathbf{z}_{\mathrm{GVL,R}})}{\gamma(\mathbf{z}_{\mathrm{GVL,E}})}. \tag{B.1}$$

Here, E denotes the extract phase and R the raffinate phase of a binary LLE between a solvent candidate X and water. P is evaluated at $T = 298.15\,\mathrm{K}$ and infinite dilution of the solute GVL is assumed.

The relative volatility α ([13]) is calculated according to

$$\alpha(T) = \frac{\gamma^{\infty}_{\mathrm{water},X}(T) \cdot p_{\mathrm{sat,water}}(T)}{\gamma^{\infty}_{\mathrm{GVL},X}(T) \cdot p_{\mathrm{sat,GVL}}(T)}, \tag{B.2}$$

where $\alpha(T)$ is evaluated at $T = 298.15\,\mathrm{K}$ with infinite dilution of water in solvent candidate X for $\gamma^{\infty}_{\mathrm{water},X}(T)$ and infinite dilution of GVL in solvent candidate X for $\gamma^{\infty}_{\mathrm{GVL},X}(T)$.

The Pearson correlation coefficient $r_{a,b}$ is defined by

$$r_{a,b} = \frac{\sum_{i=1}^{n} (a_i - \overline{a}) \left(b_i - \overline{b}\right)}{\sqrt{\sum_{i=1}^{n} (a_i - \overline{a})^2} \cdot \sqrt{\sum_{i=1}^{n} \left(b_i - \overline{b}\right)^2}}. \tag{B.3}$$

In Eq. (B.3), r is the Pearson correlation coefficient, $a = \{a_1, ..., a_i, ..., a_n\}$ and $b = \{b_1, ..., b_i, ..., b_n\}$ are data sets of possibly correlated sample values a_i and b_i and $\overline{a}$ and $\overline{b}$ are mean values of a and b respectively. The Pearson correlation coefficient $r_{a,b}$ can take on values from $r_{a,b} = -1$ to $r_{a,b} = 1$. Pearson correlation coefficients of $r_{a,b} = -1$ and $r_{a,b} = 1$ correspond to an ideal correlation of the sample values a_i and b_i. In contrast, $r_{a,b} = 0$ indicates no correlation of a_i and b_i.