



Cite this: *Digital Discovery*, 2025, 4, 3191

High-throughput application and evaluation of the COSMO-SAC model for predictions of liquid–liquid equilibria

Ivan Antolović, ^a Simon Stephan ^b and Jadran Vrabec ^{*a}

The predictive power of the COSMO-SAC activity coefficient model is rigorously put to the test using an extensive dataset of binary liquid–liquid equilibria (LLE). Two model variants, COSMO-SAC-2010 and COSMO-SAC-dsp, are evaluated across 2478 binary systems and nearly 75 000 experimental data points. They achieve a success rate exceeding 90% in detecting the occurrence of LLE, demonstrating strong qualitative performance across chemically diverse systems. In benchmark comparisons, COSMO-SAC-2010 sets the standard for nonaqueous systems, while COSMO-RS performs best for aqueous mixtures, placing the two at a broadly comparable overall level with complementary strengths. COSMO-SAC-dsp yields larger deviations but provides broader coverage, particularly for polar and hydrogen-bonding systems. Both reliably capture systematic trends across homologous series, making them effective tools for solvent screening and thermodynamic consistency analysis. A high-throughput and fully automated computational framework—integrated into the open-source package ThermoSAC—enables adaptive Gibbs energy screening, LLE tracing, and anomaly detection. This work establishes COSMO-SAC as a leading framework for predictive thermodynamics and offers reproducible benchmarks and tools for future model development, such as those based on machine learning.

Received 11th June 2025
Accepted 9th September 2025

DOI: 10.1039/d5dd00259a
rsc.li/digitaldiscovery

1 Introduction

Liquid–liquid equilibria (LLE) are critical to many industrial processes, including solvent extraction, multiphase separation, and pharmaceutical formulation. The ability to accurately predict phase behavior is crucial for designing efficient separation units, optimizing solvent selection, and ensuring process stability. While experimental determination of LLE data remains the gold standard, it is time-consuming, expensive, and often impractical for screening a large number of potential solvent systems. Thus, predictive thermodynamic models are essential for guiding research and industrial applications.

In recent years, LLE modeling has gained significant attention in various fields, including the development of sustainable separation processes and pharmaceutical drug formulation. The advent of green solvents, such as ionic liquids (IL) and deep eutectic solvents (DES), has necessitated predictive models for screening viable solvent systems efficiently.¹ Similarly, predictive modeling has been applied to aqueous sugar solutions and fruit juices to enhance separation processes in the food industry.² Moreover, LLE play a crucial role in biphasic CO₂ capture, where phase separation is used to facilitate solvent

regeneration and energy-efficient CO₂ recovery.^{3,4} In pharmaceutical applications, LLE principles govern the partitioning of drugs in aqueous two-phase systems (ATPS) and influence the miscibility of amorphous solid dispersions (ASD), impacting drug stability and bioavailability.^{5,6}

A widely used predictive model for fluid-phase equilibria is based on the continuum-solvation formalism COSMO (COnductor-like Screening MOdel),⁷ which forms the foundation of a family of COSMO-based approaches that avoid component-specific empirical parameters by using molecular surface charge distributions (σ -profiles) from quantum-chemical calculations. These distributions are used to estimate activity coefficients, enabling fully predictive phase behavior modeling. The approach was later extended to COSMO-RS (COSMO for Real Solvents)^{8–10} and COSMO-SAC (COSMO – Segment Activity Coefficient).¹¹ COSMO-SAC has been applied to vapor–liquid (VLE),^{12–14} solid–liquid (SLE),^{15–17} and LLE^{18,19} offering a unified thermodynamic framework for diverse systems.

Building upon prior research on the prediction of phase equilibria, this study extends the application of COSMO-SAC to the complex domain of LLE, representing a critical test of its generalizability. Specifically, the focus lies on the *ab initio* prediction of LLE using two COSMO-SAC variants: COSMO-SAC-2010 (no dispersion term) and COSMO-SAC-dsp (with dispersion). A large-scale evaluation is conducted on the basis of 2478

^aThermodynamics, Technische Universität Berlin, Ernst-Reuter-Platz 1, 10587 Berlin, Germany. E-mail: vrabec@tu-berlin.de

^bMolecular Thermodynamics Group (MTG), RPTU Kaiserslautern, Erwin-Schrödinger-Straße 44, 67663 Kaiserslautern, Germany



binary systems, systematically assessing the models' accuracy in predicting experimental phase behavior.

The primary objectives of this work are to (i) benchmark the predictive power of COSMO-SAC against experimental LLE data, (ii) assess whether dispersion corrections included in COSMO-SAC-dsp systematically improve predictions, (iii) develop and validate an automated workflow for high-throughput LLE screening, and (iv) provide an open-source benchmark dataset to facilitate future advancements in LLE modeling. Through large-scale computational predictions combined with rigorous validation, the study aims to refine the applicability of COSMO-SAC across diverse solvent systems and separation processes.

Beyond its methodological contributions, this work supports broader efforts to advance predictive thermodynamics for industrial applications. The large-scale evaluation highlights strengths and limitations of COSMO-SAC-2010 and COSMO-SAC-dsp for LLE modeling and informs future improvements in phase equilibrium predictions. Ultimately, the findings contribute to the development of more efficient, sustainable separation technologies and reinforce the role of computational models in solvent selection and process design.

2 Methodology

This section outlines the methodology used in this study, detailing the workflow from data acquisition to statistical analysis. The process includes compiling experimental data, determining initial values, performing LLE calculations, handling anomalies, fine-tuning results, and computing statistical metrics. Specific computational models and tools utilized in the analysis are described as well.

2.1 COSMO-SAC models

Two variants of the COSMO-SAC model were employed in this work: COSMO-SAC-2010^{11,12,20} and COSMO-SAC-dsp.²¹ A full description of this model can be found in the literature^{11,12,14,20-23} so that only a brief overview of some key aspects is provided here. The COSMO-SAC model combines quantum-chemical calculations with a statistical thermodynamic framework to predict activity coefficients and the excess Gibbs energy. Molecular surfaces are defined and described by the screening charge density σ and surface area A , which are obtained from density functional theory (DFT) calculations using a polarizable continuum model. These σ -profiles serve as an input for a statistical segment interaction model.

2.1.1 Hydrogen bonding and σ -profiles. Both COSMO-SAC-2010 and COSMO-SAC-dsp account for hydrogen bonding by categorizing molecular surface segments into non-hydrogen-bonding (NHB), hydroxyl-related hydrogen bonding (OH), and other hydrogen bonding (OT), where the latter includes segments associated with nitrogen (N), oxygen (O), and fluorine (F) atoms, as well as hydrogen atoms bonded to N or F. This segmentation is reflected in the University of Delaware (UD) database,¹³ an extension of the VT-database,^{24,25} developed in collaboration with Sandler's research group, which includes three separate σ -profiles for each hydrogen bonding type in

a single .sigma file. The University of Delaware σ -profiles, as provided by Bell *et al.*,²³ were used in this work *via* the open-source COSMO-SAC implementation available at <https://github.com/usnistgov/COSMOSAC>.

2.1.2 Dispersion term. The main difference between the two considered model variants is the treatment of the dispersive interactions. While COSMO-SAC-2010 does not explicitly include dispersion, COSMO-SAC-dsp incorporates a dispersion term. The dispersion parameter for a given molecule is determined from atomic contributions by

$$\frac{\epsilon_{\text{molecule}}}{k_B} = \frac{1}{N_{\text{atom}}} \sum_{i=1}^n \frac{\epsilon_{\text{atom},i}}{k_B}, \quad (1)$$

where $\epsilon_{\text{atom},i}$ is the dispersion parameter of atom i , k_B the Boltzmann constant, n the total number of atoms in the molecule, and N_{atom} the number of atoms with non-zero dispersion parameters (*i.e.*, for which $|\epsilon_{\text{atom},i}| > 0$). However, as noted by Hsieh *et al.*,²¹ dispersion parameters were determined for a large set of σ -profiles to demonstrate the feasibility of the approach, but not comprehensively for all available profiles. As a result, COSMO-SAC-dsp can only be applied to substances with an assigned dispersion parameter, reducing the number of applicable σ -profiles by approximately 10%. The resulting molecular dispersion parameter is then used in a combining rule to compute a binary interaction parameter, which contributes as an independent term to the total activity coefficient.

2.2 Experimental data

The predictive power of the COSMO-SAC model variants was evaluated using a comprehensive dataset compiled from the Dortmund Data Bank (DDB),²⁶ which contains experimental LLE data for binary mixtures. This dataset allows for a rigorous evaluation of model predictions against real-world measurements. To ensure meaningful statistical analysis, preprocessing was applied to retain only binary systems for which COSMO segment data (σ -profiles) are available. Additional constraints were imposed for COSMO-SAC-dsp, which requires valid dispersion parameters. The following subsections detail the dataset selection, the categorization of chemical substances, and the classification of LLE systems.

2.2.1 Dataset selection and processing. Based on the compiled experimental LLE dataset, a systematic selection and preprocessing step ensured valid COSMO-SAC model comparisons. The dataset includes 6153 unique binary systems, 120 476

Table 1 Summary of experimental LLE data and σ -profile matches. Of the 2262 profiles, 1076 match substances in the dataset—demanding profiles for both components reduces this to 933 (COSMO-SAC-2010) or 870 when dispersion parameters are also required (COSMO-SAC-dsp)

	Systems	Data points	Substances
Experimental data	6153	120 476	2462
COSMO-SAC-2010	2478	74 296	933
COSMO-SAC-dsp	2258	71 280	870



data points, and 2462 substances. To ensure consistency in model evaluation, only systems with available σ -profiles for both components were retained for the COSMO-SAC calculations. A summary of the dataset statistics is provided in Table 1.

In total, 2262 σ -profiles are present in the UD database. Among the 2462 substances compiled in the experimental dataset, 1076 have a matching σ -profile. However, only 933 can form binary systems where valid σ -profiles are available for both components, resulting in 2478 systems and 74 296 data points for evaluation.

For COSMO-SAC-dsp, an additional constraint was imposed, requiring σ -profiles to have a valid dispersion parameter. However, 237 out of the 2262 σ -profiles had a missing dispersion parameter, as the value $\varepsilon_{\text{molecule}}$ defined in eqn (1) was absent. This led to a reduction by 220 systems. Consequently, the COSMO-SAC-dsp dataset comprises 2258 binary systems with 71 280 data points formed by 870 unique substances.

Most experimental LLE data (70.1%) were measured at ambient pressure, either indicating values near 1 bar or cases with absent pressure information, for which ambient pressure was assumed. The remaining 29.9% span elevated pressures from 1.2 bar to 12 000 bar, with the majority below 2000 bar. The temperature ranges from 87.6 K to 694.6 K and exhibits a pronounced clustering around standard laboratory conditions, with 51.0% of data concentrated in the interquartile range of 293.1 K to 344.3 K.

The COSMO-SAC activity coefficient model solely depends on temperature and mole fraction, with no explicit pressure dependence. As a result, all model predictions correspond to ambient pressure conditions. Although activity coefficients in the liquid phase are often considered pressure-invariant due to low compressibility, elevated pressures can still affect phase behavior. Therefore, high-pressure experimental data were omitted during quantitative comparisons.

2.2.2 Categorization of chemical substances. Beyond dataset selection, categorizing substances by chemical nature is essential, as the predictive accuracy varies across functional groups. Each substance was assigned to a chemical family based on the scheme by Fingerhut *et al.*,¹⁴ enabling a systematic evaluation of model performance and insight into how molecular interactions influence its predictive power. It groups substances based on their dominant functional properties, arranging them in increasing order of polarity. This organization allows for the identification of trends in model performance across different chemical families. Representative examples of these families, along with their respective number of members, are presented in Table 2.

2.2.3 Classification of LLE types. To enable a systematic evaluation of LLE, a classification based on phase behavior was performed for both experimental data and predicted LLE curves. This classification is essential for understanding trends in phase separation and for refining the statistical analysis of temperature deviations. While the experimental dataset inherently distinguishes between two liquid phases (L1 and L2), it does not explicitly indicate whether a system is subject to an upper critical solution temperature (UCST) or a lower critical solution temperature (LCST) regime. Since this distinction is

Table 2 Classification of chemical families with representative examples, arranged by increasing polarity

Family	Members	Examples
Gases	4	Carbon dioxide, hydrogen sulfide, sulfur dioxide
Multifunctionals	158	Tetrachloromethane, furfural, vinyl acetate
Other nitrogens	20	Nitrobenzene, hydrazine, nitric acid
Alkanes	129	Cyclohexane, biphenyl, <i>n</i> -undecane
Alkenes	52	Cyclohexene, ethylene, 1-pentene
Alkynes	8	Phenylacetylene, ethyne, 1-hexyne
Aromatics	59	Benzene, phenol, furan
Carbonates	4	Ethylene carbonate, propylene carbonate
Epoxies	2	1,2-Propylene oxide, 2,2-dimethyloxirane
Esters	86	Ethyl acetate, dimethyl adipate, amyl formate
Halogenated hydrocarbons	77	Ethyl iodide, chloroform, dichloromethane
Halogens	2	Bromine, iodine
Ethers	43	Diethyl ether, tetrahydrofuran, ethoxybenzene
Peroxy (no acids)	2	1-Methyl-1-phenylethyl hydroperoxide
Acids	55	Formic acid, hydrogen iodide, lauric acid
Anhydrides	2	Acetic anhydride, maleic anhydride
Amines	45	Aniline, ethylenediamine, piperidine
Carbonyls	56	Acetone, 2-butanone, acrolein
Thiols	10	Ethanethiol, 2-propanethiol, hexylmercaptan
Thioethers	3	Diethyl sulfide, dimethyl sulfide, dibutylsulfide
Alcohols	91	2-Propanol, cyclopentanol, <i>n</i> -tridecanol
Amides	11	Acetamide, acetanilide, formamide
(Iso)nitriles	11	Acetonitrile, 1,4-dicyanobutane, benzylcyanide
Sulfoxides and sulfonyls	2	Dimethyl sulfoxide, sulfolane
Water	1	Water
Total	933	

crucial for adequate interpretation, a manual classification was conducted by visually inspecting and annotating all binary systems.

The following classification scheme applies specifically to the experimental data, distinguishing LLE systems based on their phase separation behavior:

- UCST: phase separation occurs upon cooling, meaning that the homogeneous phase is stable at higher temperatures, but undergoes demixing as the temperature decreases.
- LCST: phase separation occurs upon heating, where the homogeneous phase is stable at lower temperatures, but demixes as the temperature increases.
- Closed-loop: the system exhibits both UCST and LCST behavior, forming a phase separation loop with a miscibility gap in an intermediate temperature range.
- Other: unusual patterns that do not conform to classical UCST or LCST behavior, including island-type, hourglass-shaped, or diagonally aligned phase separation regions.



Table 3 Classification of experimental LLE systems based on visual inspection

LLE type	Number	Percentage
NoType	905	36.5%
WithType	1573	63.5%
-UCST	1445	91.9%
-LCST	63	4.0%
-Closed-loop	40	2.5%
-Other ^a	25	1.6%

^a Other = island-type, hourglass-shaped, and diagonally aligned LLE types.

- NoType: insufficient data prevents reliable classification.

Most classified experimental systems exhibit UCST behavior (~92%), while LCST and closed-loop systems form smaller fractions. This predominance aligns with expectations, as many known LLE systems exhibit temperature-dependent phase behavior favoring UCST-type separation. Table 3 summarizes the classification results of the experimental data.

This classification provides a framework for subsequent statistical analyses. Specifically, the interpretation of temperature deviations depends on whether a system follows UCST, LCST, or closed-loop behavior. Systems with a closed-loop phase behavior, for instance, may exhibit non-monotonic temperature deviations that require specialized analytical treatment, as discussed in Section S4 of the SI.

2.3 Initial value detection

Unlike VLE and SLE, where pure component properties such as vapor pressure or melting temperature can often serve as reference points, LLE have no such fixed anchor points. Instead, the LLE curve is detached from the pure component boundaries and must be located within the temperature-mole fraction ($T - x$) plane at a given pressure. As a result, generating suitable initial values is an essential part of the calculation process.

Beyond the lack of anchor points, LLE calculations also present challenges in solving the equilibrium conditions. Unlike other phase equilibrium types, LLE calculations require evaluating activity coefficients for each component in all phases. For a binary system, this results in four separate evaluations—compared to just one for SLE and two for VLE. This means that all four activity coefficients must be considered simultaneously, making LLE calculations more demanding. The fundamental equilibrium conditions for each phase equilibrium type are given by

$$\text{VLE: } x_i^L \gamma_i^L = \frac{y_i \cdot \varphi_i^V}{p_{0i}^{\text{LV}} \cdot \varphi_{0i}^{\text{LV}} \cdot \Pi_{0i}^{\text{LV}}}, \quad (2)$$

$$\text{SLE: } x_i^L \gamma_i^L = \exp\left(-\frac{\Delta_{\text{fus}}g_i(T)}{RT}\right), \quad (3)$$

$$\text{LLE: } x_i^L \gamma_i^L = x_i^{L1} \gamma_i^{L1} = x_i^{L2} \gamma_i^{L2}, \quad (4)$$

where x_i^L and x_i^{L1}, x_i^{L2} represent the mole fractions of component i in the liquid phase(s), γ_i^L , γ_i^{L1} , and γ_i^{L2} denote the corresponding activity coefficients, y_i is the vapor-phase mole fraction, and φ_i^V and φ_{0i}^{LV} are the fugacity coefficients of the component in the mixture and as a pure substance, respectively. The term p_{0i}^{LV} refers to the pure component vapor pressure, while Π_{0i}^{LV} denotes the Poynting correction. For SLE, $\Delta_{\text{fus}}g_i(T)$ represents the Gibbs energy change upon fusion at the temperature T , with R being the gas constant.

As shown in eqn (2)–(4), each phase equilibrium type imposes different modeling requirements when expressed in terms of activity coefficients. The VLE includes fugacity coefficients of the vapor phase, while SLE depend on the Gibbs energy of fusion to describe the equilibrium between the solid and liquid phases. In contrast, LLE rely solely on activity coefficients, without the need for experimental reference data or an equation of state. This makes LLE structurally distinct in modeling and highly sensitive to the choice of the activity coefficient model. At the same time, LLE are an ideal benchmark for evaluating thermodynamic models, as any discrepancies directly reflect their ability to capture liquid-phase interactions.

Due to its sole dependence on the activity coefficient model, LLE are highly sensitive to model configurations. Unlike VLE and SLE, which show minor variations upon parameter changes, LLE predictions can shift drastically. This was evidenced in our preceding work,⁶ where altering the combinatorial term—switching from Staverman–Guggenheim to free volume or removing it—led to an “entropic catastrophe”. While SLE remained relatively stable upon such changes, LLE showed extreme variations, highlighting strong model sensitivity.

2.3.1 Forward screening. Initially, experimental data were used as starting values for LLE calculations, but in most cases, neither temperature nor composition were suitable. Considering 2478 binary systems with about 75 000 data points further underscored this challenge, highlighting the need for a more systematic approach. Instead of relying on experimental data, the focus shifted entirely to the model itself to generate reliable initial values for LLE calculations.

To achieve this, a structured screening approach was developed to identify initial values solely from the thermodynamic model. By systematically evaluating the Gibbs energy of mixing Δg_{mix} across a range of temperatures and a fixed number of mole fractions, a comprehensive search was conducted in the $T - x$ plane, independent of system-specific experimental data, as shown in Fig. 1. The Gibbs energy of mixing was determined by

$$\frac{\Delta g_{\text{mix}}}{RT} = \sum_i x_i \ln a_i = \sum_i x_i \ln (x_i \gamma_i), \quad (5)$$

where x_i denotes the mole fraction of component i and a_i corresponds to its activity.

The temperature range for screening was set between 0 K and 1000 K. To improve computational efficiency, an adaptive temperature step size was implemented. For temperatures above 200 K, a step size of 40 K was applied, while in the intermediate range between 100 K and 200 K, a finer step size of 20 K was used. Below 100 K, where phase separation is expected



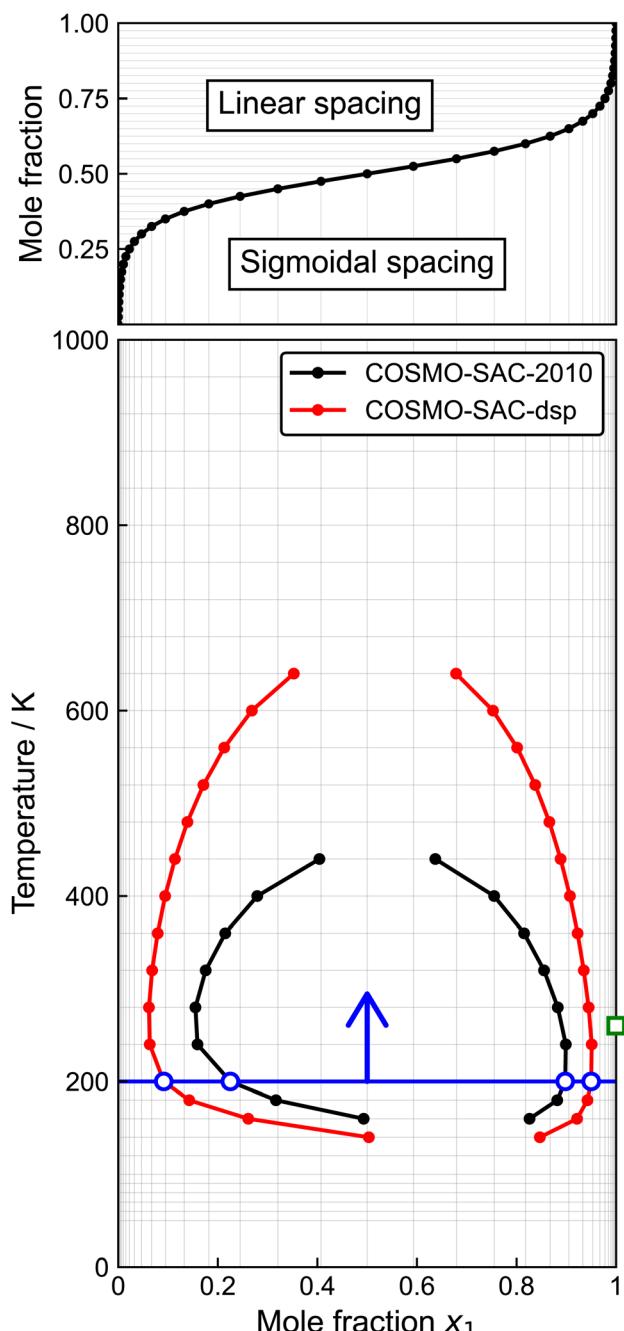


Fig. 1 Forward screening process for identifying initial values during LLE tracing, illustrated for the system ethylene glycol + 2,5-dimethyltetrahydrofuran. Screening progresses from low to high temperatures, systematically evaluating the Gibbs energy of mixing Δg_{mix} across the $T - x$ space. Mole fraction spacing followed a sigmoidal transformation to improve resolution near the infinite dilution limits. The temperature range was adaptively sampled between 0 K to 1000 K, with finer steps at lower temperatures where phase separation is more pronounced. For clarity, a reduced number of grid lines is shown; actual screenings discretized the mole fraction space with 201 to 1001 points depending on system complexity. The green square marks the melting point of ethylene glycol.

to be more pronounced, the step size was further refined to 10 K to enhance resolution.

It is important to note that the predicted LLE curves may extend below the melting point of one or both components, indicating metastable states with respect to SLE. For example, pure ethylene glycol solidifies near 260 K (see Fig. 1). While SLE are not modeled in this study, the associated thermodynamic constraint is acknowledged and has been addressed previously in the context of COSMO-SAC-based SLE modeling.⁶

To improve LLE detection, mole fraction spacing was adapted in addition to temperature. While initial scans used uniform spacing, analysis of 2478 binary systems showed that higher resolution near the infinite dilution limits significantly improved detection and initial value quality. A sigmoidal transformation based on a logistic function was applied to redistribute the originally equidistant grid accordingly. The mapping function is given by

$$S(x_i) = \frac{\sigma(\lambda(x_i - 0.5)) - \sigma(-\lambda/2)}{1 - 2\sigma(-\lambda/2)}, \quad (6)$$

where $\sigma(t) = 1/(1 + \exp(-t))$ is the logistic sigmoid function, the inflection parameter was set to $\lambda = 25$ to concentrate sample points near the infinite dilution limits. This refined sampling strategy not only improved accuracy in critical phase separation regions, but also reduced the need for additional scans, making the detection process more efficient.

Since most experimental systems exhibit a UCST (91.9%), screening was performed from low to high temperatures, denoted as forward screening. Larger miscibility gaps typically occur at lower temperatures and yield more pronounced Gibbs energy curves, facilitating LLE detection by providing clearer separation signals.

Conversely, screening from high to low temperatures often first encounters the UCST region, which is already challenging to detect, even when its exact location is known. Furthermore, initial values found near the UCST often lead to uncertainties about whether true equilibrium conditions have been identified, as calculations in this region tend to scatter and require refinement (see Section S1 of the SI). To ensure stability and robustness in detecting LLE, it is therefore preferable to begin the search at temperatures far away from critical solution temperatures.

However, reverse screening from high to low temperatures was conducted as well to identify and mitigate “early scan errors”. This step ensures that no relevant phase separation regions are overlooked during forward screening. Once a potential LLE binodal was detected, screening was stopped to avoid unnecessary computational effort, as this initial value served as a starting point for subsequent LLE tracing, which is described in Section 2.4.

From the Gibbs energy of mixing curve, the binodal was determined using the alternating tangents method proposed by von Solms *et al.*²⁷ First, the spinodal points were identified, and the derivative was approximated using central differences. Starting from one spinodal point, a tangent was constructed toward the opposite spinodal point, as shown in Fig. 2. This approach was found to be both robust and reliable. While the

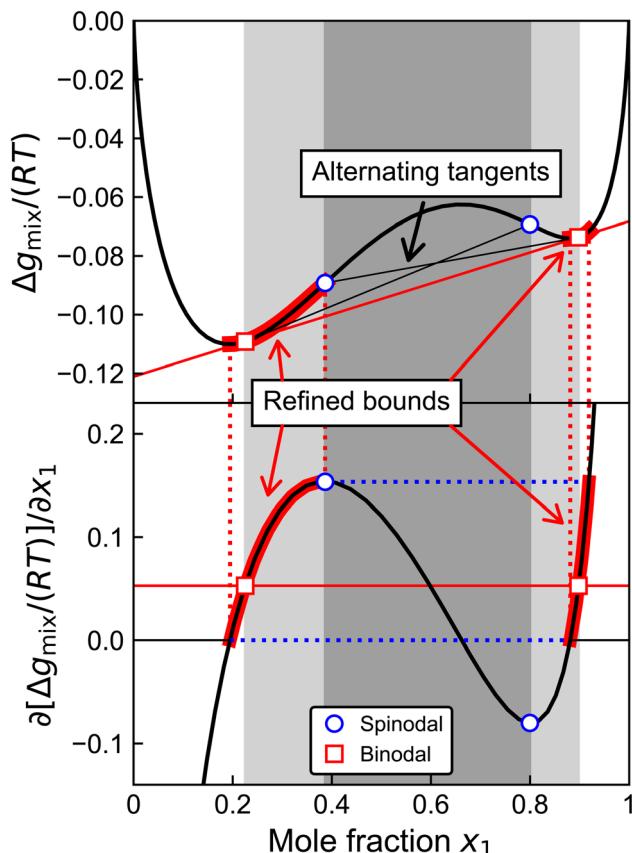


Fig. 2 Identification of the binodal with the alternating tangents method. The Gibbs energy of mixing $\Delta g_{\text{mix}}/(RT)$ was analyzed to locate the spinodals, which served as starting points for tangent construction. Alternating tangents were drawn between the spinodals to determine equilibrium compositions, refining the binodal. Additional boundary refinement, based on local minima of Δg_{mix} , improved accuracy, particularly in cases with multiple LLE regions.

original method defines bounds between the spinodal points and the pure component edges, an improvement was introduced by incorporating derivative information. Since the tangent must have equal slopes at the equilibrium points, the boundaries were further refined to lie between the spinodal and local minima of the Gibbs energy of mixing curve. This refinement is particularly useful when handling multiple LLE regions at a single temperature.

Ultimately, the initial value screening provides two key pieces of information: (1) whether an LLE exists for the given system, and (2) its temperature and mole fraction ($T - x$) coordinates. These data form the basis for further calculations, ensuring that numerical solvers start from well-defined initial conditions. By combining systematic screening with bidirectional evaluation, the accuracy and robustness of LLE detection are significantly improved.

2.3.2 Reverse screening. While the forward screening approach provides a systematic route to detect LLE regions, it was primarily optimized for identifying the dominant miscibility gap at lower temperatures. Some systems, however, exhibit detached island-type regions—phase separation pockets

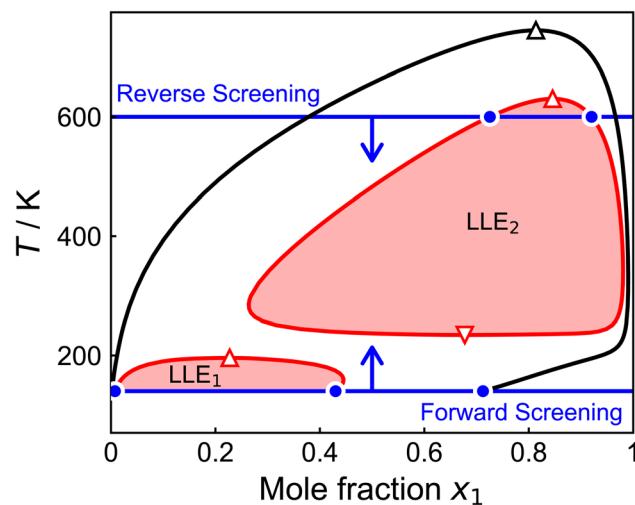


Fig. 3 The system water + valeronitrile shows two disconnected, island-type LLE regions with COSMO-SAC-dsp (red), while COSMO-SAC-2010 predicts a single continuous LLE (black). Forward screening detected only LLE₁; reverse screening revealed the second island LLE₂. Blue markers indicate the first detected binodal in each direction.

that appear at higher temperatures, disconnected from the primary miscibility gap. Since screening was halted once an LLE region was detected, these isolated regions could remain unnoticed, necessitating an additional step to ensure comprehensive phase equilibrium detection.

To address this, a reverse screening procedure was introduced, screening from high to low temperatures. This method served as a safeguard against early scan errors, where the algorithm terminates after identifying the first LLE region, potentially missing additional phase separation regions. Such errors were particularly evident in systems where the detected LLE did not align with experimental data, suggesting the presence of another, previously undetected binodal at higher temperatures. By reversing the screening direction, these additional LLE regions—otherwise overlooked—were successfully identified.

Reverse screening revealed previously undetected phase separation in four water-containing systems. These island-type LLE regions appeared only with COSMO-SAC-dsp, while COSMO-SAC-2010 predicted a single, continuous region. A good example is water + valeronitrile, as shown in Fig. 3, where COSMO-SAC-dsp predicts two distinct LLE regions, while COSMO-SAC-2010 yields just one. Forward screening became trapped in the smaller LLE₁, missing the larger LLE₂ at higher temperatures. Reverse screening successfully identified LLE₂, resolving early scan errors and improving agreement with experimental data.

2.4 LLE tracing

Once initial values T_0 , x_0^{L1} , and x_0^{L2} were found, the first binodal was obtained by solving eqn (4). Subsequent points were computed by advancing a temperature step ΔT and using the previous binodal as the initial guess. This tracing approach is conceptually similar to isochoric tracing,^{28,29} where phase



boundaries are constructed starting from a pure component and marching into the mixture, as demonstrated by Bell and Deiters³⁰ and further explored in later works.^{31,32}

As the tracing proceeds in temperature direction, LLE curves can terminate in two distinct ways: (a) at a critical point, where the compositions of the two liquid phases become identical, or (b) at a vapor–liquid–liquid equilibrium (VLLE), where the LLE binodal intersects a VLE envelope, forming a three-phase region. While this study does not explicitly resolve VLLE due to the absence of vapor-phase fugacities, binodal termination behavior was still considered during anomaly detection and binodal classification.

However, selecting an appropriate step size ΔT is not trivial. If too large, it can produce rough or inaccurate curves, particularly near the critical solution temperatures, where the slope $(\partial T/\partial x_i)_p$ of the binodal approaches zero. In extreme cases, the UCST or LCST may be missed due to overshooting. Conversely, a very small step size ensures high resolution, but significantly increases computational cost due to the large number of required steps.

Manual step size adjustment is feasible for a small number of systems, but impractical for tackling 2478 systems and two model variants, amounting to approximately 5000 different LLE curves. A constant step size may perform well in steep regions but remains inadequate near the UCST. To ensure robustness, reproducibility, and computational efficiency, an adaptive sampling method with a dynamically adjusted ΔT was required.

2.4.1 Adaptive sampling. To improve upon the constant step size approach, several adaptive methods were tested. The most straightforward approach was to couple the step size to the width of the miscibility gap, defined as $\Delta x \equiv |x_i^{L2} - x_i^{L1}|$. By making ΔT proportional to that width, the step size naturally decreases near the UCST, where Δx approaches zero, leading to a higher resolution in critical regions. Conversely, in steep regions of the LLE curve, a larger step size is permitted

$$\Delta T \propto \Delta x^n. \quad (7)$$

The exponent n was introduced to refine resolution near the UCST when necessary ($n > 1$), as seen in a polymer amorphous–amorphous phase separation study.⁶ However, in this work, the exponent was kept at unity, $n = 1$. This simple adjustment significantly improved results over the constant step-size method, as illustrated in Fig. 4a and c.

Several adaptive methods were explored, including those inspired by mechanical motion (velocity, acceleration), geometric concepts (arc length), and Taylor approximations. While some offered theoretical advantages, they often suffered from numerical instability or lacked robustness across all systems. Further details, including equations and performance insights, are given in Table S1 of the SI. Ultimately, the simple miscibility-gap scaling according to eqn (7) proved to be most reliable and broadly effective.

2.4.2 Predictive initial guess. Beyond the temperature step size ΔT , the efficiency and accuracy of LLE curve calculations depend on the choice of an initial guess for subsequent iterations. A naive approach—using the previous result as the next initial guess—becomes problematic near the UCST, where the slope $\Delta x' = \partial(\Delta x)/\partial T$ is large and the true solution may deviate significantly from the previous result. In regions with pronounced curvature (large second derivative $\Delta x'' = \partial^2(\Delta x)/\partial T^2$), this approach can lead to poor predictions, requiring more iterations for convergence that increase computational cost.

To improve efficiency, a dynamically adjusted initial guess was implemented using a Taylor expansion. Although the step-size methods discussed earlier are formulated in terms of the miscibility gap width Δx , the actual predictive update applies to the vector of mole fractions, $\mathbf{x}_1^{\text{LLE}} = (x_1^{\text{L1}}, x_1^{\text{L2}})$. Nevertheless, for consistency with the previous equations, we retain the notation Δx in the predictive expression

$$\Delta x(T + \Delta T) \approx \Delta x + \Delta x' \Delta T + \frac{\Delta x''}{2} \Delta T^2 + \mathcal{O}(\Delta T^3), \quad (8)$$

where the term $\mathcal{O}(\Delta T^3)$ accounts for higher-order contributions, which were neglected. This predictive adjustment improves robustness, particularly in challenging LLE regions,

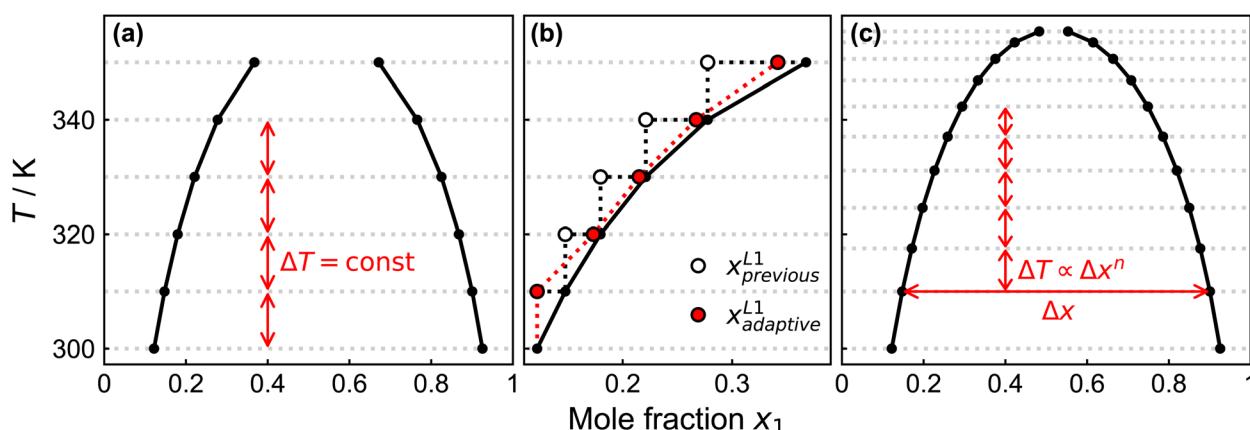


Fig. 4 Comparison of adaptive sampling methods for LLE tracing. (a) Constant step size yields rough curves near the UCST. (b) Adaptive initial guesses improve accuracy and convergence. (c) Step size scaled with Δx enhances resolution and smoothness near the UCST without compromising efficiency.



as shown in Fig. 4b. By accounting for both the first and second order derivative terms, the initial guess better aligns with the true solution, reducing the number of required iterations and thereby improving computational efficiency.

2.4.3 Resolving anomalies. Initial screening typically resulted in well-defined LLE curves, though some systems showed anomalies that required further analysis. Common cases—such as single data point results from premature equilibrium detection—were resolved by rescreening or refined initial guesses. More persistent issues, including temperature jumps, widening miscibility gaps, and irregular binodal shapes, often stemmed from overly large step sizes. These were addressed with targeted corrections; further details are provided in Section S4 of the SI.

The most frequent anomaly observed was a sudden bend of the LLE curve with rising temperature, inconsistent with meaningful phase separation. This irregularity, termed here as a *bifurcation anomaly*, involves an abrupt slope change in one curve branch—typically the one farther from the pure component boundary. While the widening miscibility gap initially suggests a UCST, the distorted curvature signals a breakdown of equilibrium. An analysis of the Gibbs energy of mixing reveals the root cause: instead of two, these systems exhibit four spinodal points, allowing for two distinct LLE regions, each with its own LCST. As temperature rises, the curves intersect at a point where the inner binodal points coincide, marking the onset of the anomaly. Beyond this point, the tangents connecting inner and outer binodal points intersect, and the system transitions into a single LLE.

This transition is illustrated in Fig. 5, where inset (a) shows the Gibbs energy of mixing below the intersection point, featuring two well-separated miscibility gaps. These represent two independent LLE regions governed by their respective

LCST. As temperature rises, the inner binodal points converge, merging the two curves. Inset (b) captures the behavior above the intersection, where only a single LLE remains and the inner binodal tangents become unstable, as described by Novák *et al.*³³ Notably, four spinodal points do not always yield dual LLE curves unless the tangents connecting binodal pairs are true tangents, not secants—a condition requiring the inner and outer points to align consistently without producing physically unrealistic crossings.

Such complex phase behavior is not merely theoretical. The binary system methylisopropyl ketone + water exhibits two biphasic regions at different pressures and even three coexisting liquid phases (LLLE) at intermediate pressure, as shown experimentally by Steiner and Schadow.³⁴ COSMO-SAC predicts this behavior from first principles, supporting the plausibility of bifurcation anomalies and demonstrating its capability to capture complex multiphase equilibria. The alternating tangents method²⁷ was extended to handle multiple spinodal regions, enabling robust treatment and preventing misinterpretation.

2.5 Statistical evaluation

Following the computational modeling of LLE with the COSMO-SAC framework, a quantitative assessment of the predictive accuracy was performed by comparing calculated and experimental values of key thermodynamic properties. The primary variables of interest are temperature T and mole fraction x_i . Another variable that lends itself for statistical evaluation in the context of LLE is the distribution coefficient K_i , which condenses the entire phase equilibrium information of a system into a single value

$$K_i = \frac{x_i^{L1}}{x_i^{L2}} = \frac{\gamma_i^{L2}}{\gamma_i^{L1}}. \quad (9)$$

Here, x_i^{L1} and x_i^{L2} are the mole fractions of component i in the two liquid phases, and γ_i^{L1} , γ_i^{L2} are the corresponding activity coefficients. To quantify deviations from experiment, the average deviation (AD) and average absolute deviation (AAD) were computed

$$AD_X = \frac{1}{N} \sum_{k=1}^N (X_k^{\text{calc}} - X_k^{\text{exp}}), \quad (10)$$

$$AAD_X = \frac{1}{N} \sum_{k=1}^N |X_k^{\text{calc}} - X_k^{\text{exp}}|. \quad (11)$$

Therein, X serves as a placeholder for the variables x_1 , T , or K_1 , while N is the number of experimental data points. The AAD reflects the average magnitude of deviation, while the AD captures systematic bias, with positive and negative values indicating over- or underprediction, respectively.

2.5.1 Projection alignment for statistical evaluation.

Unlike mole fraction deviation, which primarily requires phase consistency between experimental and predicted values, temperature deviation demands additional structural considerations. Mole fraction deviations are calculated by matching each experimental point to the corresponding phase (L1 or L2)

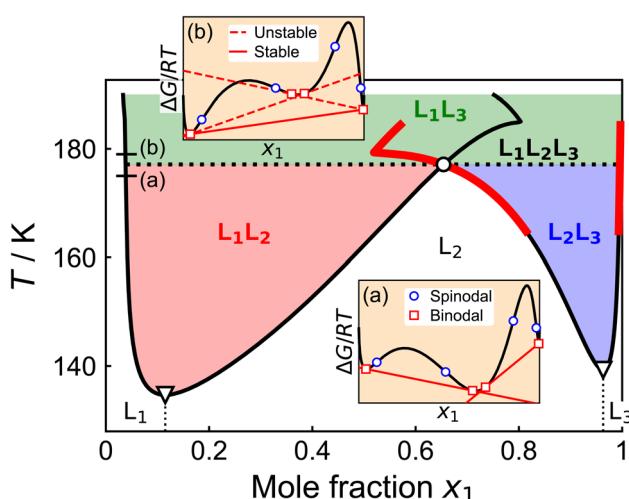


Fig. 5 Illustration of the bifurcation anomaly of the water + hexanol system predicted by COSMO-SAC-2010. The LLE curve initially widens but bends sharply at higher temperatures (bold red line). Gibbs energy analysis reveals two distinct LLE regions with separate LCST that merge into a single LLE. Panel (a) shows two LLE regions below the merging point; panel (b) shows one stable LLE above.



on the predicted binodal. Temperature deviations, by contrast, require distinguishing between different regions of the binodal curve—particularly in systems with closed-loop or non-UCST behavior. In such cases, experimental points must be projected vertically onto the corresponding section of the calculated curve (see Fig. S4 of the SI).

3 Results

COSMO-SAC-2010 and COSMO-SAC-dsp were benchmarked for their ability to detect LLE, classify phase behavior, and quantitatively predict experimental binodals. Qualitative prediction success is first assessed, followed by a detailed classification of LLE types and critical phenomena. Quantitative accuracy is then analyzed through deviations in terms of mole fraction, temperature, and distribution coefficient. Finally, systematic trends of homologous series are examined, highlighting the models' capabilities in solvent screening applications.

3.1 Preliminary assessment of LLE predictions

The first step in evaluating the model performance is to examine their ability to qualitatively predict whether LLE exist for a given binary system. While this comparison does not yet evaluate the accuracy of the predicted phase boundaries, it merely provides an overview of how the models differ in their general capability to capture LLE occurrence.

Fig. 6 summarizes the success rates of both model variants under two conditions: the full dataset comprising 2478 binary systems with available experimental data and σ -profiles, and a reduced dataset limited to 2258 systems with assigned dispersion parameters $\varepsilon_{\text{molecule}}$. In the full dataset, COSMO-SAC-2010 identified LLE in 2155 systems and COSMO-SAC-dsp

in 2075, with combined predictions capturing 2240 systems (90.4%). This number is likely to be higher if dispersion parameter data were complete.

Because COSMO-SAC-dsp applies only to systems with valid $\varepsilon_{\text{molecule}}$ values, its predictions are restricted to a subset of 2258 binary systems. Within this set, COSMO-SAC-2010 and COSMO-SAC-dsp predict LLE in 1994 and 2075 cases, respectively, with a combined success rate of 92.1%. Importantly, COSMO-SAC does not predict LLE by default; each result emerges from the model's underlying thermodynamics. The high success rates thus reflect genuine predictive performance, not an artifact of the method.

The results in Fig. 6 reveal several key trends: COSMO-SAC-2010 achieves success rates of 87.0% and 88.3% in the full and constrained datasets, respectively. COSMO-SAC-dsp reaches 83.7% in the full dataset, with an estimated increase to 90.5% if dispersion parameters were available for all σ -profiles. Despite its broader coverage, COSMO-SAC-dsp does not supersede the 2010 variant: it misses LLE in four systems captured by COSMO-SAC-2010, while uniquely identifying 85 systems not detected by the latter. These differences underscore the complementary nature of the two model variants.

3.2 Classification and characterization of LLE types

To evaluate the diversity of phase behavior predicted by the COSMO-SAC model, a classification scheme was established based on structural features in computed phase diagrams. Unlike experimental data, which reveal a broad range of LLE types, COSMO-SAC predictions consistently yield at least one UCST per system—except in rare diverging cases.

Fig. 7 shows typical phase behaviors predicted by COSMO-SAC. Most common are (a) UCST and (b) closed loop LLE with both UCST and LCST. Other topologies include (c) hourglass shapes, (d) droplet-like boundaries, and (e) valley-shaped gaps. Some systems exhibit (f) emerging or (g) fully detached immiscibility regions. More complex cases include (h) double LCST, (i) LCST with an island, and (j) two UCST and two LCST forming distinct, closed loops. The most intricate case is (j), with two UCST and two LCST. Unlike bifurcation anomalies, both form distinct closed loops, resembling a fused dual LLE system.

Further deviations from classical behavior appear in panels (k) and (l). One system exhibits a diverging LLE, with boundaries extending beyond the screening range. Another reveals a chromosome-like topology where the solver passed a pinch point and uncovered a second stable solution. The lower branch, indicated by a dashed line, is interpreted as a metastable artifact, reflecting the occasional stabilization of spurious LLE solutions in systems with island-type complexity. Interestingly, none of the COSMO-SAC predictions display a purely LCST-driven demixing—a significant departure from experimental observations. Since LCST behavior is generally associated with entropy-driven phase separation,^{35,36} this discrepancy suggests that the COSMO-SAC models may over-emphasize enthalpic interactions, potentially due to inherent modeling assumptions.

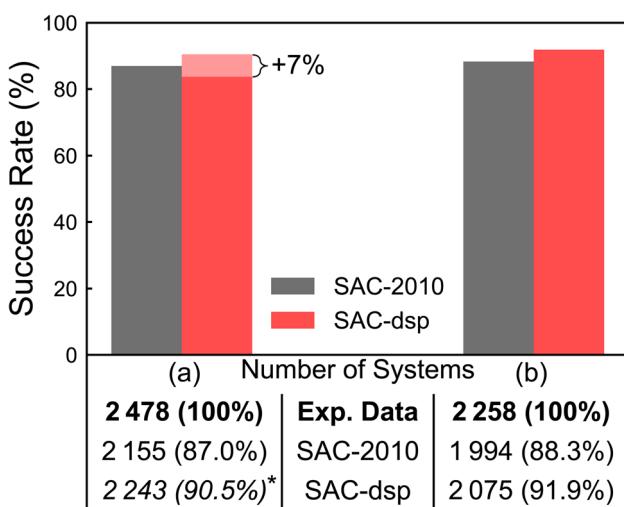


Fig. 6 LLE prediction success rates of COSMO-SAC-2010 and COSMO-SAC-dsp. (a) Full dataset of 2478 binary systems with available experimental data and σ -profiles. (b) Reduced dataset of 2258 binary systems where valid dispersion parameters $\varepsilon_{\text{molecule}}$ are also available. The asterisk (*) indicates an estimated 7% increase in predicted LLE occurrences, if dispersion parameters would be available for all relevant substances.



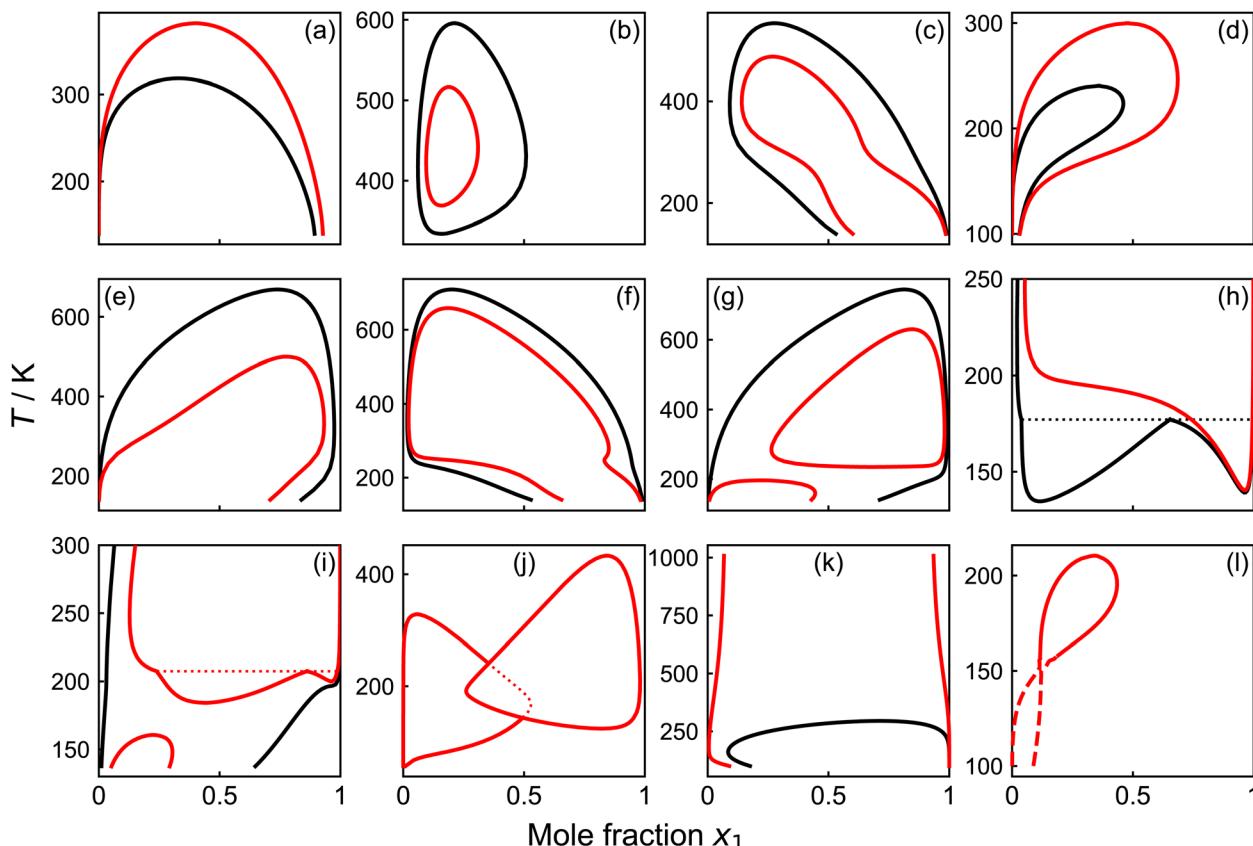


Fig. 7 Qualitatively different LLE types predicted by COSMO-SAC-2010 (black) and COSMO-SAC-dsp (red). The classification distinguishes between traditional behaviors, such as UCST (a) and closed-loop (b), and more complex structures, including hourglass (c), raindrop (d), valley (e), and island (g). No purely LCST-driven separations were predicted, but multiple UCST or UCST-LCST combinations. Systems shown are listed in Table S3 of the SI.

3.3 Mutually exclusive binodals and critical point distribution

An intriguing observation emerges when comparing the LLE curves predicted by COSMO-SAC-2010 and COSMO-SAC-dsp: their binodal envelopes are always distinct and either fully envelop or are enveloped by the other, without mutual intersection. This strict nesting behavior, evident in Fig. 7, permits a significant simplification: each LLE curve can be represented by a single scalar value, namely the UCST. Since the binodals never intersect, a higher UCST directly implies a broader miscibility gap, making it a robust metric for comparing the two model variants. This also reveals systematic tendencies—if one model overestimates the miscibility gap, the other necessarily does so to a greater extent.

This trend is visualized in Fig. 8, which shows the UCST distribution and associated critical mole fractions. For clarity, mole fractions were mirrored such that COSMO-SAC-2010 data appear in the left half ($x_i \in [0, 0.5]$) and COSMO-SAC-dsp in the right half ($x_i \in [0.5, 1]$). This transformation— $x_i \rightarrow \min(x_i, 1 - x_i)$ or $x_i \rightarrow \max(x_i, 1 - x_i)$, depending on the variant—preserves physical validity while enhancing interpretability and symmetry. It enables a clearer analysis of clustering and systematic behavior across both models.

Most predicted UCST lie between 200 K to 500 K, with critical mole fractions clustering around $x_i = 0.25$ to 0.75 . A noticeable high-temperature cluster above 700 K, composed almost entirely of aqueous systems, appears in both model variants (see Fig. S5 of the SI). At lower temperatures, critical points tend to shift toward composition edges, suggesting broader asymmetry. Despite the lack of thermodynamic anchor points and the model's sole reliance on γ_i , nearly all predicted UCST lie within the experimentally sampled range of 87.6 K to 694.65 K, underscoring the predictive power of COSMO-SAC.

Statistically, COSMO-SAC-dsp predicts a higher UCST than COSMO-SAC-2010 for 79.6% of all systems, indicating a systematic overestimation of the miscibility gap width. In the subset of cases where COSMO-SAC-dsp yields a lower UCST, 87.7% involve aqueous mixtures, suggesting that dispersion corrections disproportionately affect strongly polar or hydrogen-bonded systems.

3.3.1 Absence of UCST. About 9% of COSMO-SAC-dsp predictions (192 out of 2075 systems) exhibit no UCST within the scanned range up to 2000 K, indicating diverging LLE behavior. This anomaly, shown in Fig. 7k, is exclusive to COSMO-SAC-dsp and strongly linked to carbon dioxide (54.7%), nitromethane (39.6%), and nitroethane (5.2%). These compounds share polar double-bonded oxygen groups—such



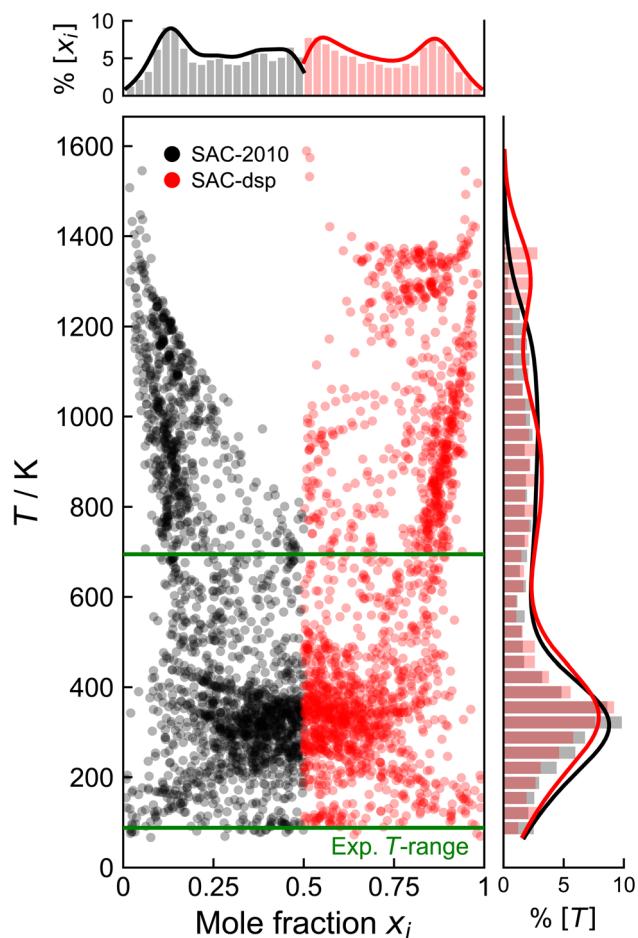


Fig. 8 Predicted UCST and corresponding mole fractions for all systems where both COSMO-SAC-2010 (black) and COSMO-SAC-dsp (red) yield LLE. To enhance visual clarity and distinguish between model variants, mole fractions were mirrored: COSMO-SAC-2010 is shown on the left half ($x_i \in [0, 0.5]$) and COSMO-SAC-dsp on the right half ($x_i \in [0.5, 1]$). Histograms display the distributions of UCST and critical mole fractions, respectively. Green horizontal lines indicate bounds where experimental LLE data are available.

as the linear carbonyl-like $\text{C}(=\text{O})=\text{O}$ structure of CO_2 and the nitro group $\text{N}(=\text{O})\text{O}^-$ —which may promote long-range directional interactions. Further details are provided in Section S5 of the SI.

3.4 Quantitative assessment of predictive accuracy

To provide a more detailed quantitative assessment of model accuracy, the AAD_x of the mole fraction was evaluated across all systems, grouped by chemical family combinations. Each binary system was assigned to a pair of chemical families, enabling a dense and informative matrix representation of family–family interactions, as shown in Fig. 9. The focus on mole fraction deviations is motivated by the fact that LLE are inherently defined at a fixed temperature, and mole fractions directly describe the composition of the coexisting liquid phases, offering a natural basis for comparison. This focus enables a detailed evaluation while retaining a concise discussion.

Deviations in other quantities—such as T , K_1 , x_1^{L1} , and x_2^{L1} —are provided in the SI.

To ensure a robust quantitative comparison, additional data exclusion steps were applied beyond that for the qualitative assessment. Specifically, all experimental data points measured at pressures above 10 bar were excluded, following the procedure of Fingerhut *et al.*¹⁴ In their VLE study, this threshold was chosen to approximate ideal gas behavior. In the context of LLE, although the liquid phase is often treated as incompressible, pressure effects cannot be fully neglected at elevated pressures. Experimental LLE data frequently showed steep or discontinuous curve segments, which likely belong to remote isobars and are not directly comparable to ambient-pressure predictions.

3.4.1 Pairwise evaluation of chemical family interactions.

Fig. 9 uses a dual-circle grid to compare the AAD_x of both model variants for each family pair. In total, 107 family–family combinations were evaluated. Among these, 87 are shared by both model variants, while 20 are only represented by COSMO-SAC-dsp. This does not necessarily imply that COSMO-SAC-2010 failed to predict LLE in those cases; rather, it may reflect the absence of corresponding experimental data points at comparable x_1 values, which are required for evaluating AAD_x under the applied projection scheme.

Among the 87 directly comparable chemical family pairs, COSMO-SAC-2010 performs better in 44 cases (*i.e.* 50.6% of all pairs), COSMO-SAC-dsp in 20 cases (23.0%), with near-identical accuracy (<1% difference) for the remaining 23 cases (26.4%). The most notable improvements from dispersion were found for amines + aromatics (ΔAAD_x reduced from 50.5% to 21.9%), alkanes + halogenated hydrocarbons (ΔAAD_x : -12.7%), and acids + aromatics (ΔAAD_x : -11.6%). Aqueous systems also benefit notably, *e.g.*, water + epoxies (ΔAAD_x : -11.0%) and water + (iso)nitriles (ΔAAD_x : -8.6%).

Conversely, the highest AAD_x for COSMO-SAC-dsp occurs for esters + multifunctionals (48.2%) and gas-containing pairs, due to carbon dioxide. The largest deteriorations appear for other nitrogens with ethers, alkenes, alcohols, or alkanes (up to ΔAAD_x : +21.6%), and for alkanes + anhydrides. Still, over a quarter of family pairs yield comparable AAD_x for both models, especially in weakly interacting systems, such as alcohols + gases or alkanes + ethers. Best performance is seen for water with alkynes, alkanes, alkenes, or halogenated hydrocarbons, and for (iso)nitriles or amides with aromatics.

3.4.2 Individual family contributions to predictive error.

Fig. 10 shows the AAD_x for each individual chemical family. Across most families, COSMO-SAC-dsp exhibits systematically larger AAD_x compared to its 2010 counterpart, reaffirming earlier observations of a general overprediction of the miscibility gap width upon inclusion of the dispersion term. A notable exception is observed for epoxies, where COSMO-SAC-dsp yields a smaller AAD_x . However, this result is based on a single data point *versus* 23 data points in COSMO-SAC-2010, rendering the comparison statistically unreliable.

The largest individual AAD_x values occur for the “Gases” and “OtherNitrogens” families, though closer inspection reveals that these families are represented by only a few components after exclusion based on pressure, dispersion parameters, and

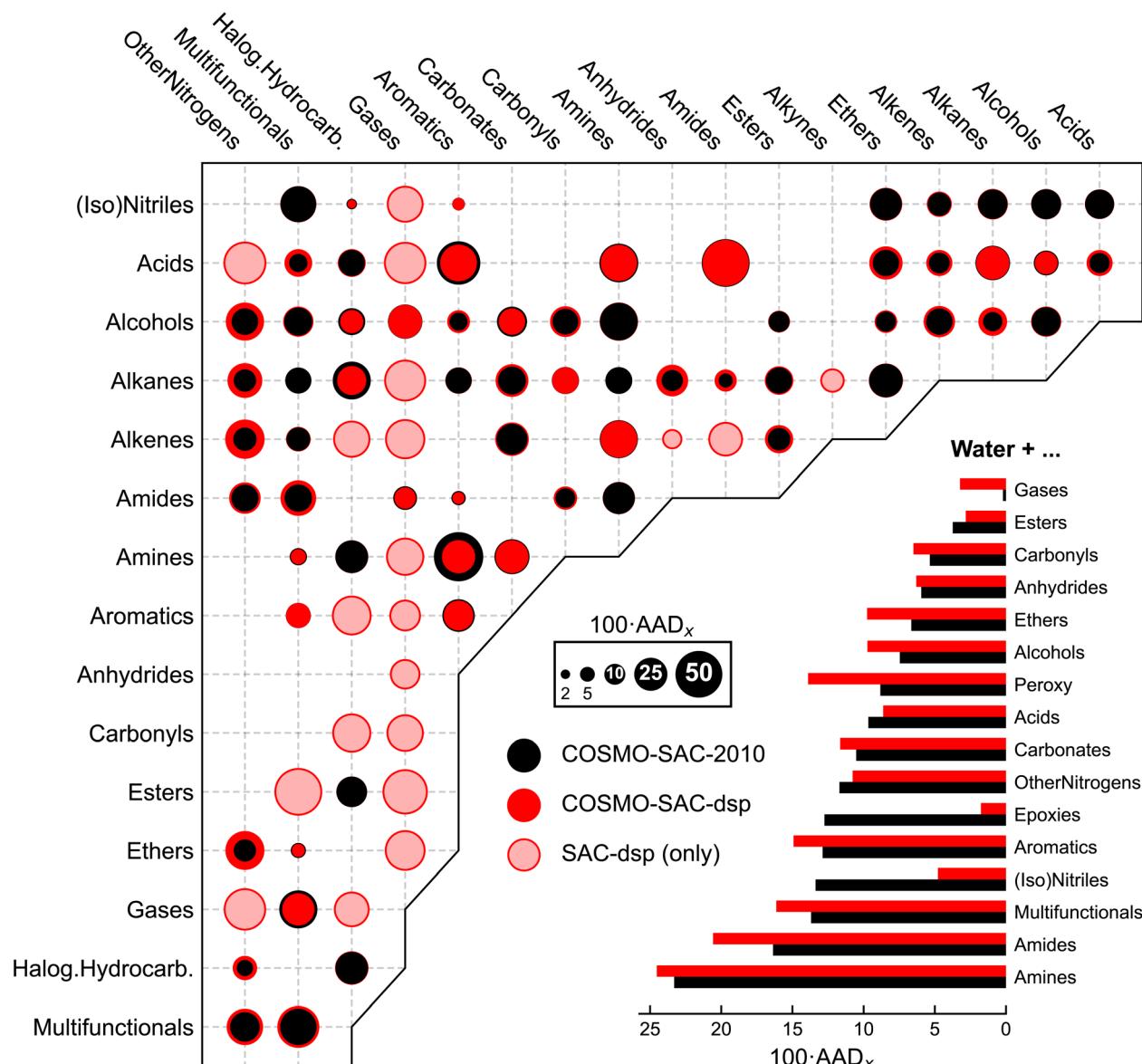


Fig. 9 Grid of all occurring chemical family pairs, with circle areas indicating the $100 \cdot AAD_x$ values for COSMO-SAC-2010 and COSMO-SAC-dsp. The smaller circle is always plotted on top to visualize which model variant performs better. Pairs involving water are treated separately and shown below the main grid with a dedicated bar plot.

data availability. In the case of “Gases”, only carbon dioxide remains, while the “OtherNitrogens” family is dominated by nitromethane and nitroethane—precisely the systems where COSMO-SAC-dsp predicts diverging LLE behavior, leading to disproportionately large deviations. In contrast, chemically more balanced families, such as acids and esters, exhibit similar AAD_x values across both model variants. Overall, the AAD_x amounts to 9.9% for COSMO-SAC-2010 and 13.0% for COSMO-SAC-dsp, with the difference largely driven by these few problematic systems.

This result is remarkable when viewed in the broader context of previous studies. The total AAD_x for LLE predictions aligns closely with previous solubility studies applying COSMO-SAC to pharmaceutical systems (12.6%),⁶ despite the absence of

external anchor points such as the melting temperature used in SLE. Furthermore, activity coefficient deviations of COSMO-SAC can span a wide range—from 0% near the pure component limit to as much as 300% at infinite dilution.¹⁴ Given that LLE is determined solely by γ_i across the entire composition space, this level of accuracy is notable. By contrast, the lower AAD_x typically observed in VLE can be partially attributed to the presence of a reference point, such as a boiling point, which constrains both temperature and composition predictions. Furthermore, the evaluation by Fingerhut *et al.*¹⁴ was limited to vapor-phase mole fractions, whereas LLE requires accurate predictions for both coexisting liquid phases.

3.4.3 Overall model accuracy and systematic deviations. Table 4 summarizes the statistical performance of both



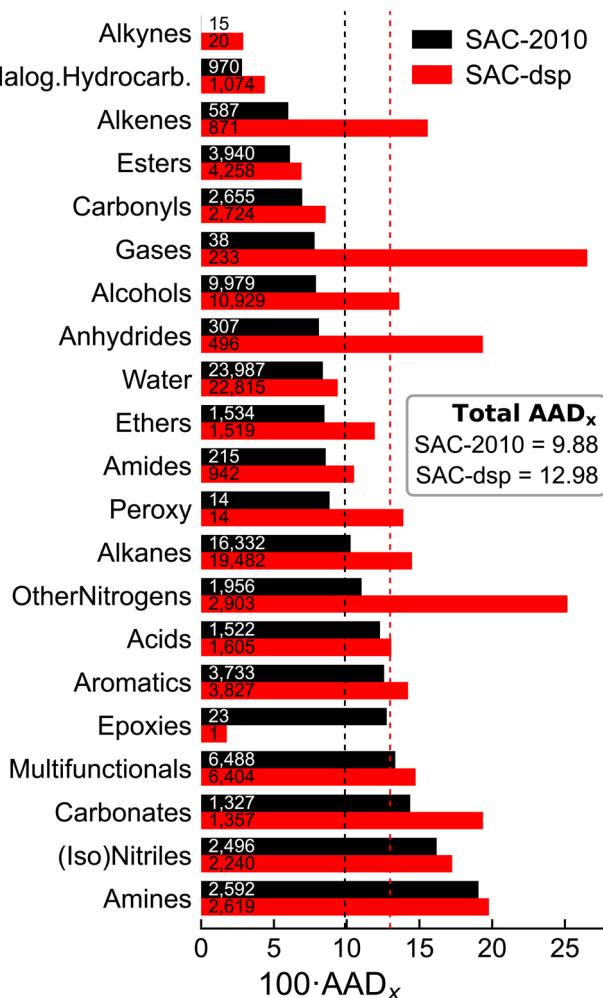


Fig. 10 AAD_x in terms of mole fraction for each chemical family, comparing COSMO-SAC-2010 (black) and COSMO-SAC-dsp (red). Bar lengths scale with $100 \cdot AAD_x$; numerical labels indicate the number of experimental data points per family. Dashed lines mark the total AAD_x of the two model variants. COSMO-SAC-dsp generally shows larger deviations, mainly driven by outliers such as CO_2 ("Gases") and polar nitrogen compounds ("OtherNitrogens").

COSMO-SAC model variants across all evaluated systems. It reports the AD, AAD, and number of contributing data points N_{eff} for the key variables, including temperature T , overall composition x_1 , phase-specific mole fractions x_1^{L1} , x_1^{L2} , and distribution coefficient K_1 . While the total AAD_x offers a general measure, a phase-specific analysis reveals systematic trends. For example, both COSMO-SAC variants consistently underestimate x_1^{L1} and overestimate x_1^{L2} , resulting in a near-zero overall AD for x_1 due to cancellation—yet clearly indicating a tendency to overpredict the miscibility gap width. This is also reflected by the negative AD_K values of K_1 .

The contributing data points are well balanced not only between phases, but also between temperature and mole fraction for COSMO-SAC-2010, ensuring an overall unbiased evaluation. COSMO-SAC-dsp generally covers more data by predicting LLE for a larger number of systems, but also exhibits more temperature mismatches. This is because diverging LLE—

Table 4 Statistical comparison of COSMO-SAC-2010 and COSMO-SAC-dsp over all evaluated LLE systems, reporting AD, AAD, and the effective number of data points N_{eff} for mole fractions x_1 , x_1^{L1} , x_1^{L2} , temperature T , and distribution coefficient K_1

	COSMO-SAC-2010			COSMO-SAC-dsp		
	AD	AAD	N_{eff}	AD	AAD	N_{eff}
x_1	0.0021	0.0988	40 449	0.0017	0.1298	43 269
T	35.85 K	81.83 K	40 444	62.56 K	99.85 K	39 069
K_1	-0.0717	0.1344	9803	-0.1031	0.1707	10 343
x_1^{L1}	-0.0587	0.1052	20 212	-0.0959	0.1369	21 659
x_1^{L2}	0.0627	0.0923	20 237	0.0995	0.1227	21 610

predicted exclusively by COSMO-SAC-dsp—lack a defined UCST, preventing meaningful comparison and thus reducing the effective sample size for AAD_T .

The AAD_T reaches approximately 82 K for COSMO-SAC-2010 and 100 K for COSMO-SAC-dsp, which may appear large. However, due to the geometry of typical LLE envelopes—flat near the UCST and steep at the binodal edges—small composition errors can produce large temperature deviations, and *vice versa*. This sensitivity is examined in Section S7 of the SI using synthetic LLE curves. The results show that minor shifts in mole fraction can cause significant temperature errors when projected onto steep binodal regions, and small temperature offsets can distort composition near the critical points. These effects highlight the need for caution when interpreting AAD in LLE contexts and suggest that additional metrics may help to provide a fuller picture of model performance.

3.5 Benchmark against other predictive models

A dedicated benchmark was performed to compare COSMO-SAC with established predictive models using a reference LLE dataset. This dataset was compiled and analyzed in the context of model comparison by Elliott *et al.* and is published in the 6th edition of the textbook *The Properties of Gases and Liquids*.³⁷ It contains 96 binary systems composed of 78 substances and a total of 12 546 individual data points, offering a solid basis for reliable benchmarking. The dataset is provided by the NIST Thermodynamics Research Center (TRC) and is accessible at <https://pgl6ed.byu.edu/>.

Table 5 compares the performance of various LLE prediction methods using the reference dataset. Unlike the AAD_x metric used throughout this work, the textbook employs the percentage average absolute logarithmic deviation (AALDS), defined as

$$\text{AALDS} \equiv \frac{1}{n} \sum_{i=1}^n \left| \ln \left(\frac{x_i^{\text{calc}}}{x_i^{\text{exp}}} \right) \right| \times 100, \quad (12)$$

where x_i is the mole fraction of the dilute component.

Among all models considered in ref. 37, COSMO-SAC outperforms COSMO-RS for nonaqueous systems, with COSMO-SAC-2010 setting the benchmark by achieving the lowest AALDS of all models. For aqueous mixtures, the picture is reversed: COSMO-RS performs best, while COSMO-SAC-2010 is



Table 5 Evaluation of methods for LLE prediction. Results for UNIFAC and COSMO-RS are adapted from Elliott *et al.*³⁷ COSMO-SAC results were calculated in this study

Method	Nonaqueous			Aqueous			Overall	
	Systems	Points	AALDS	Systems	Points	AALDS	Systems	AALDS
UNIFAC-VLE	47	4334	86.3	42	7576	62.1	89	70.9
UNIFAC-NIST/KT	42	3982	79.8	42	7576	124.6	84	109.2
UNIFAC-NIST/Mod	47	4334	49.5	41	7537	66.9	82	60.5
UNIFAC-Dortmund	43	5063	65.1	36	5962	68.2	79	66.8
COSMO-RS/FSAC2	25	3419	124.4	27	4376	61.0	52	88.8
COSMO-RS/GAMESS	40	4204	139.9	36	5905	55.3	76	90.5
COSMO-SAC-2010	34	3121	37.9	47	7772	112.6	81	91.2
COSMO-SAC-dsp	39	3805	116.7	45	7178	130.9	84	126.0

moderate and COSMO-SAC-dsp weakest. Overall, COSMO-SAC-2010 reaches a level comparable to COSMO-RS, with each showing complementary strengths. While the UNIFAC-NIST/Mod variant attains the lowest overall AALDS, this reflects extensive parameterization to experimental data, whereas COSMO-SAC and COSMO-RS are fully predictive. From this perspective, COSMO-SAC emerges as a solid, open-source alternative to COSMO-RS.

3.6 Capturing systematic behavior with COSMO-SAC

Beyond overall predictive accuracy, it is worthwhile to assess how well COSMO-SAC captures systematic trends within homologous series. This aspect is particularly relevant when absolute agreement with experimental data is less critical than understanding the relative behavior of similar systems. One representative case is shown in Fig. 11, which depicts the predicted LLE envelopes for mixtures of acetonitrile with *n*-alkanes.

Despite deviations in the exact location of the binodal curves, COSMO-SAC-2010 successfully reproduces key qualitative trends observed experimentally. In particular, it captures the systematic shift of the critical point toward higher temperature and higher acetonitrile mole fraction as the chain length

increases. This behavior reflects the rising hydrophobicity of longer alkanes and their reduced affinity for polar solvents such as acetonitrile.

Moreover, the model accurately predicts the change of the LLE envelope shape across the series: while the *n*-butane mixture exhibits an almost symmetric phase diagram, the systems with longer alkanes—such as hexadecane—show pronounced asymmetry, with the binodal curve being increasingly skewed toward pure acetonitrile. The fact that such shifts in topology are qualitatively captured without any parameter optimization, highlights the power of COSMO-SAC in predicting relative tendencies within chemical families.

From a practical perspective, this makes COSMO-SAC particularly useful for solvent screening tasks, where relative positioning of miscibility gaps may be more informative than exact critical temperatures. Applications include the selection of extraction solvents, where homologous series like alkanes, alcohols, or esters are frequently evaluated, as well as the design of phase-change solvent systems for separations or pharmaceutical formulations. In such contexts, the ability to reliably capture directional trends within chemical families—even without perfect numerical agreement—can significantly accelerate and de-risk early-stage decision making. This underscores the potential of COSMO-based models not just as quantitative predictors, but as powerful qualitative guides in chemical process design.

While only COSMO-SAC-2010 results are shown in Fig. 11 for clarity, it is worth noting that COSMO-SAC-dsp exhibits comparable predictive behavior.

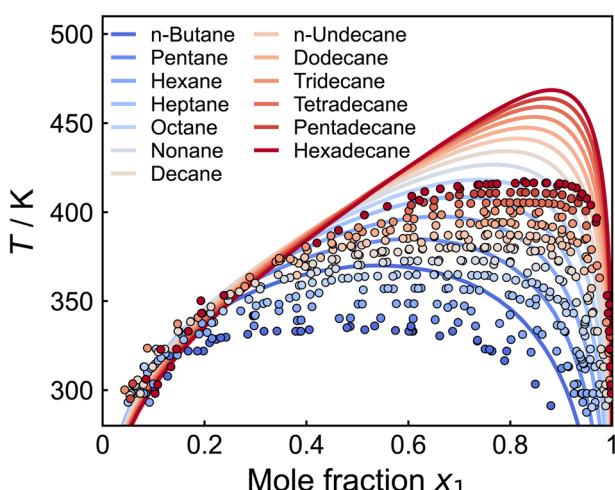


Fig. 11 Predicted LLE envelopes for mixtures of acetonitrile with *n*-alkanes (*n*-butane to *n*-hexadecane) using COSMO-SAC-2010.

3.7 Enhancing reproducibility and screening with ThermoSAC

The complete LLE workflow developed in this study is available in the open-source Python package ThermoSAC. It also integrates the SLE functionality from previous work,⁶ continuing the SLE.solubility() method from COSMOPharm and complementing it with a substantially expanded LLE.miscibility() interface. Built with a modular, object-oriented design inspired by phasepy,³⁸ ThermoSAC represents components and mixtures as class objects that interface with activity coefficient models, enabling unified workflows across SLE and LLE.

The LLE.miscibility() method encapsulates the full adaptive tracing algorithm developed in this work, including forward and reverse screening, automatic step-size control, and optional spinodal detection. A dedicated GMixScanner class further supports robust and parallelizable detection of initial LLE regions through Gibbs energy scans, with methods such as find_first_binodal() and find_all_binodal() tailored for early screening and anomaly detection. The package, including documentation and examples, is available at <https://github.com/ivanantolo/thermosac> and is intended as a reusable tool for model development, reproducibility, and educational applications.

4 Conclusion

This study presents a systematic and large-scale evaluation of the COSMO-SAC activity coefficient model for predicting LLE across diverse binary systems. Two model variants—COSMO-SAC-2010 and COSMO-SAC-dsp—are assessed over more than 2400 binary mixtures and nearly 75 000 experimental data points, providing a detailed benchmark of predictive capabilities, limitations, and applicability across chemical families and LLE types.

A central contribution of this work lies in the development of an automated, robust workflow for high-throughput LLE tracing, featuring forward and reverse screening, adaptive sampling, and anomaly detection. The framework allows for precise localization of phase boundaries and critical points, and facilitates comprehensive statistical evaluation across systems with diverse phase behaviors—including UCST, LCST, and complex bifurcating or closed-loop topologies.

The COSMO-SAC model achieves a qualitative success rate exceeding 90% in identifying LLE occurrence, demonstrating robustness across chemically diverse systems without reliance on system-specific tuning. Benchmarking against other predictive models shows that COSMO-SAC outperforms COSMO-RS for nonaqueous systems, with COSMO-SAC-2010 setting the benchmark. For aqueous mixtures the trend is reversed, with COSMO-RS performing best, while COSMO-SAC-2010 is moderate and COSMO-SAC-dsp weakest. Overall, COSMO-SAC-2010 is comparable to COSMO-RS, with complementary strengths, whereas the UNIFAC-NIST/Mod variant attains the lowest total AALDS through parameterization to experimental data.

Both COSMO-SAC variants effectively capture systematic trends within homologous series, correctly predicting relative shifts in critical temperature and asymmetry of the miscibility gap. This capability is valuable for solvent screening and process design, where relative performance within chemical families often outweighs the need for exact numeric agreement. Quantitatively, COSMO-SAC-2010 achieves a total AAD_x of 9.9% and AAD_T of 81.8 K, confirming it as the more precise variant overall. COSMO-SAC-dsp provides broader coverage, though with slightly larger deviations (13.0% AAD_x , 99.9 K AAD_T).

In summary, this work delineates the performance landscape of COSMO-SAC for LLE prediction and provides an open-source infrastructure with benchmark data to support future

model development. The present insights advance the use of COSMO-based methods in solvent screening and process design, with implications for separation technologies, green chemistry, and pharmaceutical applications. Ongoing research into refined interaction terms, extended segment descriptors, and hybrid modeling strategies—as exemplified by developments in openCOSMO-RS, which incorporates London-type dispersion potentials^{39–42}—promises to further enhance the reliability and reach of first-principles phase equilibrium predictions.

Conflicts of interest

There are no conflicts to declare.

Data availability

All data supporting the findings of this study are available at <https://github.com/ivanantolo/thermosac>. This includes the full LLE tracing results for all evaluated binary systems, initial values generated by the GmixScanner, critical point data (UCST and LCST), mappings of substances to chemical families, statistical evaluations of model accuracy, and the open-source ThermoSAC Python package implementing the complete workflow. The specific version of the ThermoSAC package used in this study has been archived at <https://doi.org/10.5281/zenodo.17073717>.

Supplementary information: methodological details on LLE tracing refinements, adaptive sampling strategies, and anomaly corrections, together with additional figures and statistical analyses. It further includes critical point distributions and special system classes (aqueous and diverging systems) as well as complete system lists for figures in the main text. See DOI: <https://doi.org/10.1039/d5dd00259a>.

Acknowledgements

I. A. and J. V. gratefully acknowledge financial support by Deutsche Forschungsgemeinschaft (DFG) under grant no. VR 6/16. The authors thank Dr Jürgen Rarey for fruitful discussions. The authors also thank J. Richard Elliott for helpful feedback.

References

- 1 J. Yang, Z. Hou, G. Wen, P. Cui, Y. Wang and J. Gao, A brief review of the prediction of liquid–liquid equilibrium of ternary systems containing ionic liquids by the COSMO-SAC model, *J. Solution Chem.*, 2019, **48**, 1547–1563.
- 2 L. T. Paese, R. L. Spengler, R. d. P. Soares and P. B. Staudt, Predicting phase equilibrium of aqueous sugar solutions and industrial juices using COSMO-SAC, *J. Food Eng.*, 2020, **274**, 109836.
- 3 Y.-C. Hung, C.-M. Hsieh, H. Machida, S.-T. Lin and Y. Shimoyama, Towards design of phase separation solvent for CO₂ capture using COSMO-SAC model, *J. Mol. Liq.*, 2021, **336**, 116229.



4 Y.-C. Hung, C.-M. Hsieh, H. Machida, S.-T. Lin and Y. Shimoyama, Unveiling the mechanism of CO₂-driven phase change in amine+ water+ glycol ether ternary mixture, *J. Taiwan Inst. Chem. Eng.*, 2022, **131**, 104143.

5 F. Ghaffari, M. Khorsandi, H. Shekaari and M. T. Zafarani-Moattar, Liquid-liquid equilibrium measurements and computational study of salt-polymer aqueous two phase system for extraction of analgesic drugs, *Sci. Rep.*, 2022, **12**, 13848.

6 I. Antolović, J. Vrabec and M. Klajmon, COSMOPharm: Drug-Polymer Compatibility of Pharmaceutical Amorphous Solid Dispersions from COSMO-SAC, *Mol. Pharm.*, 2024, **21**, 4395–4415.

7 A. Klamt and G. Schüürmann, COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient, *J. Chem. Soc., Perkin Trans.*, 1993, **2**, 799–805.

8 A. Klamt, Conductor-Like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena, *J. Phys. Chem.*, 1995, **99**, 2224–2235.

9 A. Klamt, V. Jonas, T. Bürger and J. C. W. Lohrenz, Refinement and Parametrization of COSMO-RS, *J. Phys. Chem. A*, 1998, **102**, 5074–5085.

10 A. Klamt, *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam, 2005.

11 S. T. Lin and S. I. Sandler, A Priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model, *Ind. Eng. Chem. Res.*, 2002, **41**, 899–913.

12 C. M. Hsieh, S. I. Sandler and S. T. Lin, Improvements of COSMO-SAC for Vapor-Liquid and Liquid-Liquid Equilibrium Predictions, *Fluid Phase Equilib.*, 2010, **297**, 90–97.

13 R. C. Xiong, S. I. Sandler and R. I. Burnett, An Improvement to COSMO-SAC for Predicting Thermodynamic Properties, *Ind. Eng. Chem. Res.*, 2014, **53**, 8265–8278.

14 R. Fingerhut, W. L. Chen, A. Schedemann, W. Cordes, J. Rarey, C. M. Hsieh, J. Vrabec and S. T. Lin, Comprehensive Assessment of COSMO-SAC Models for Predictions of Fluid-Phase Equilibria, *Ind. Eng. Chem. Res.*, 2017, **56**, 9868–9884.

15 D. Peng, A. Alhadid and M. Minceva, Assessment of COSMO-SAC predictions for solid-liquid equilibrium in binary eutectic systems, *Ind. Eng. Chem. Res.*, 2022, **61**, 13256–13264.

16 B. Bouillot, S. Teychene and B. Biscans, An evaluation of COSMO-SAC model and its evolutions for the prediction of drug-like molecule solubility: part 1, *Ind. Eng. Chem. Res.*, 2013, **52**, 9276–9284.

17 S. Z. Mahmoudabadi and G. Pazuki, Investigation of COSMO-SAC model for solubility and cocrystal formation of pharmaceutical compounds, *Sci. Rep.*, 2020, **10**, 19879.

18 M. R. Shah and G. D. Yadav, Prediction of liquid-liquid equilibria of (aromatic+ aliphatic+ ionic liquid) systems using the Cosmo-SAC model, *J. Chem. Thermodyn.*, 2012, **49**, 62–69.

19 C. L. da Silveira and N. P. G. Salau, The UNIFAC-LLE and COSMO-SAC ternary aqueous LLE calculations, *Fluid Phase Equilib.*, 2019, **501**, 112278.

20 S. Wang, S. I. Sandler and C.-C. Chen, Refinement of COSMO-SAC and the Applications, *Ind. Eng. Chem. Res.*, 2007, **46**, 7275–7288.

21 C.-M. Hsieh, S.-T. Lin and J. Vrabec, Considering the Dispersive Interactions in the COSMO-SAC Model for More Accurate Predictions of Fluid Phase Behavior, *Fluid Phase Equilib.*, 2014, **367**, 109–116.

22 W.-L. Chen, C.-M. Hsieh, L. Yang, C.-C. Hsu and S.-T. Lin, A Critical Evaluation on the Performance of COSMO-SAC Models for Vapor-Liquid and Liquid-Liquid Equilibrium Predictions Based on Different Quantum Chemical Calculations, *Ind. Eng. Chem. Res.*, 2016, **55**, 9312–9322.

23 I. H. Bell, E. Mickoleit, C. M. Hsieh, S. T. Lin, J. Vrabec, C. Breitkopf and A. A. Jager, Benchmark Open-Source Implementation of COSMO-SAC, *J. Chem. Theory Comput.*, 2020, **16**, 2635–2646.

24 E. Mullins, R. Oldland, Y. A. Liu, S. Wang, S. I. Sandler, C. C. Chen, M. Zwolak and K. C. Seavey, Sigma-Profile Database for Using COSMO-Based Thermodynamic Methods, *Ind. Eng. Chem. Res.*, 2006, **45**, 4389–4415.

25 E. Mullins, Y. A. Liu, A. Ghaderi and S. D. Fast, Sigma Profile Database for Predicting Solid Solubility in Pure and Mixed Solvent Mixtures for Organic Pharmacological Compounds with COSMO-Based Thermodynamic Methods, *Ind. Eng. Chem. Res.*, 2008, **47**, 1707–1725.

26 Dortmund Data Bank, Version 2023. <http://www.ddbst.com>, Version 2023 released on April 13, 2023. Accessed: May 23, 2023.

27 N. von Solms, I. A. Kouskoumvekaki, T. Lindvig, M. L. Michelsen and G. M. Kontogeorgis, A Novel Approach to Liquid-Liquid Equilibrium in Polymer Systems with Application to Simplified PC-SAFT, *Fluid Phase Equilib.*, 2004, **222**, 87–93.

28 S. E. Quiñones-Cisneros and U. K. Deiters, An efficient algorithm for the calculation of phase envelopes of fluid mixtures, *Fluid Phase Equilib.*, 2012, **329**, 22–31.

29 U. K. Deiters and T. Kraska, *High-Pressure Fluid Phase Equilibria – Phenomenology and Computation*, Elsevier, Amsterdam, 1st edn, 2012, vol. 2.

30 I. H. Bell and U. K. Deiters, On the construction of binary mixture p-x and T-x diagrams from isochoric thermodynamics, *AIChE J.*, 2018, **64**, 2745–2757.

31 U. K. Deiters and I. H. Bell, Calculation of phase envelopes of fluid mixtures through parametric marching, *AIChE J.*, 2019, **65**, e16730.

32 I. H. Bell, U. K. Deiters and A. Jäger, Algorithm to Identify Vapor-Liquid-Liquid Equilibria of Binary Mixtures from Vapor-Liquid Equilibria, *Ind. Eng. Chem. Res.*, 2022, **61**, 2592–2599.

33 J. P. Novák, J. Matouš and J. Pick, *Liquid-Liquid Equilibria*, Academia, Prague, 1st edn, 1987.

34 R. Steiner and E. Schadow, Visuelle und dekametrische Bestimmung von Phasengleichgewichten nichtmischbarer



Flüssigkeiten unter Hochdruck, *Z. Phys. Chem.*, 1969, **63**, 297–311.

35 I. C. Sanchez and M. T. Stone, in *Polymer Blends Volume 1: Formulation*, ed. D. R. Paul and C. B. Bucknall, John Wiley & Sons, New York, 2000, pp 15–54.

36 S. Nastyshyn, Y. Stetsyshyn, J. Raczkowska, Y. Nastishin, Y. Melnyk, Y. Panchenko and A. Budkowski, Temperature-responsive polymer brush coatings for advanced biomedical applications, *Polymers*, 2022, **14**, 4245.

37 J. R. Elliott, V. Diky, T. A. Knotts IV and W. V. Wilding, *The Properties of Gases and Liquids*, McGraw-Hill Education, 6th edn, 2023.

38 G. Chaparro and A. Mejía, Phasepy: A Python based framework for fluid phase equilibria and interfacial properties computation, *J. Comput. Chem.*, 2020, **41**, 2504–2526.

39 T. Gerlach, S. Müller, A. G. de Castilla and I. Smirnova, An open source COSMO-RS implementation and parameterization supporting the efficient implementation of multiple segment descriptors, *Fluid Phase Equilib.*, 2022, **560**, 113472.

40 F. London, The General Theory of Molecular Forces, *Trans. Faraday Soc.*, 1937, **33**, 8b–26.

41 D. Grigorash, S. Müller, P. Paricaud, E. H. Stenby, I. Smirnova and W. Yan, A comprehensive approach to incorporating intermolecular dispersion into the openCOSMO-RS model. Part 1. Halocarbons, *Chem. Eng. Sci.*, 2025, **309**, 121425.

42 D. Grigorash, S. Müller, E. Heid, F. Neese, D. Liakos, C. Riplinger, M. García-Ratés, P. Paricaud, E. H. Stenby, I. Smirnova and W. Yan, A comprehensive approach to incorporating intermolecular dispersion into the openCOSMO-RS model. Part 2: atomic polarizabilities, *Chem. Eng. Sci.*, 2025, **319**, 122170.