RIO

PUC

Dissertação de Mestrado

Thermodynamic modelling of vapor – liquid equilibria of strongly non-ideal systems: the cyclo-hexane – ethanol binary.

Jonathan Melo de Azevedo

Pontifícia Universidade Católica do Rio de Janeiro
Centro Técnico Científico
Departamento de Engenharia Química e de Materiais

Rio de Janeiro, 16 de setembro de 2025

Pontifícia
Universidade
Católica do
Rio de Janeiro

Dissertação de Mestrado

Thermodynamic modelling of vapor – liquid equilibria of strongly non-ideal systems: the cyclo-hexane – ethanol binary.

Jonathan Melo de Azevedo

Orientação: Professor Rogério Navarro Correia de Siqueira
Coorientação: Professor Jadran Vrabec

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia Química, de Materiais e Processos Ambientais, do Departamento de Engenharia Química e de Materiais da PUC-Rio.

Rio de Janeiro, 16 de setembro de 2025

Pontifícia
Universidade
Católica do
Rio de Janeiro

Thermodynamic modelling of vapor – liquid equilibria of strongly non-ideal systems: the cyclo-hexane – ethanol binary.

Jonathan Melo de Azevedo

**Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia
Química, de Materiais e Processos Ambientais da PUC-Rio.
Aprovada pela Comissão Examinadora abaixo:**

Professor Rogério Navarro Correia de Siqueira
Orientador
Departamento de Engenharia Química e de Materiais - PUC-Rio

Professor Jadran Vrabec
Co-Orientador
Technische Universität Berlin (TU-Berlin)

Professor Roberto Ribeiro de Avillez
Departamento de Engenharia Química e de Materiais - PUC-Rio

Professor Fernando Luiz Pellegrini Pessoa
Centro Universitário SENAI CIMATEC

Rio de Janeiro, 16 de setembro de 2025

Pontifícia
Universidade
Católica do
Rio de Janeiro

Jonathan Melo de Azevedo

Graduado em Engenharia Química pela Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio, em 2023. Áreas de interesse: Termodinâmica Computacional e Engenharia Química.

Ficha Catalográfica

Azevedo, Jonathan Melo de

Thermodynamic modelling of vapor – liquid equilibria of strongly non-ideal systems: the cyclo-hexane – ethanol binary / Jonathan Melo de Azevedo ; advisor: Rogério Navarro Correia de Siqueira ; co-advisor: Jadran Vrabec. – 2025.
117 f. : il. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Química e de Materiais, 2025.
Inclui bibliografia

1. Engenharia Química e de Materiais – Teses. 2. Ciclohexano-etanol. 3. Equilíbrio líquido-vapor. 4. UNIFAC-Do. 5. Peng-Robinson. I. Siqueira, Rogério Navarro Correia de. II. Vrabec, Jadran. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Química e de Materiais. IV. Título.

CDD: 620.11

Acknowledgments

I would like to express my deepest gratitude to my advisor, Prof. Rogério N. C. de Siqueira, and my co-advisor, Prof. Jadran Vrabec for their guidance and support throughout the course of this research. Their expertise, insights and feedback were fundamental in the elaboration of this dissertation.

I would like to thank the Pontifical Catholic University of Rio de Janeiro and all its faculty members who provided invaluable guidance throughout my journey in Chemical Engineering.

Finally, and most importantly, I would like to thank my family and friends. To my parents, Roselane and Jonas, for all their unconditional love and belief in me. To God, for placing such wonderful people in my life. I would not have made it here without Him.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

Melo de Azevedo, Jonathan; Navarro C. de S., Rogério (Advisor); Vrabec, Jadran (Co-Advisor). **Thermodynamic modelling of vapor – liquid equilibria of strongly non-ideal systems: the cyclo-hexane – ethanol binary.** Rio de Janeiro, 2025. 117p. Dissertação de Mestrado – Departamento de Engenharia Química e de Materiais, Pontifícia Universidade Católica do Rio de Janeiro.

Vapor–liquid equilibria (VLE) are fundamental to distillation, one of the most widely employed separation techniques in chemical and pharmaceutical industries. The presence of asymmetrical molecules in shape and size, combined with polar functional groups, results in strong intermolecular interactions, leading to significant deviations from ideal mixture behavior. Consequently, accurately modeling such non-ideal systems, particularly those forming azeotropes, remains a demanding task that requires robust thermodynamic models and reliable parameter estimation strategies. The binary cyclohexane – ethanol is an example of such system, with high importance for the ethanol chemical industry, where its separation remains a key step in the production of anhydrous ethanol. In this study, group contribution parameters for the UNIFAC model were optimized using the Nelder–Mead algorithm (downhill simplex method), implemented in MATLAB, based on literature isothermal VLE data (γ–φ formalism) in the temperature range of 278–338 K. Next, isobaric phase equilibria (40–150 kPa) were evaluated and compared with experimental data, with particular attention to the performance difference relative to the UNIFAC-Do (Dortmund) model with parameters obtained from the DDB database. The influence of the initial guess in the optimization process was further investigated, as well as the effect of alternative distributions of activity coefficient data obtained from COSMO-SAC calculations (either concentrated in the cyclohexane-rich region, in the ethanol-rich region, or uniformly spaced across the composition range). Finally, a hybrid optimization strategy combining successive optimization cycles was proposed and tested, which proved to be the most balanced among the investigated approaches, since it surpassed the isothermal-fitted parameters under isobaric conditions and outperformed the UNIFAC-Do model in most isothermal cases, while also achieving a complementary performance in reproducing activity coefficients. During all computations, the Peng–Robinson equation of state with classical van der Waals mixing rules was employed to calculate fugacity coefficients in the vapor

phase. The comparison between optimized UNIFAC parameters and UNIFAC-Do results highlighted the trade-off between isothermal and isobaric predictions, as well as the ability of the hybrid routine to improve overall consistency. Regarding azeotrope prediction, the proposed methodology enabled a satisfactory reproduction of both temperature and composition at 101.3 kPa. It is worth noting that the formalism adopted here has not been previously applied to the cyclohexane–ethanol system, thus reinforcing the novelty and relevance of this contribution.

Keywords

Computational Thermodynamics; Peng – Robinson; UNIFAC; Buble pressure.

Resumo

Melo de Azevedo, Jonathan; Navarro C. de S., Rogério (Orientador); Vrabec, Jadran (Coorientador). **Modelagem termodinâmica do equilíbrio líquido-vapor de sistemas fortemente não ideais: o binário ciclo-hexano – etanol**. Rio de Janeiro, 2025. 117p. Dissertação de Mestrado – Departamento de Engenharia Química e de Materiais, Pontifícia Universidade Católica do Rio de Janeiro.

Os equilíbrios líquido-vapor (ELV) são fundamentais para a destilação, uma das técnicas de separação mais amplamente utilizadas nas indústrias química e farmacêutica. A presença de moléculas assimétricas em forma e tamanho, combinada com grupos funcionais polares, resulta em fortes interações intermoleculares, levando a desvios significativos em relação ao comportamento de misturas ideais. Consequentemente, modelar com precisão tais sistemas não ideais, particularmente aqueles que formam azeótropos, continua sendo uma tarefa desafiadora que requer modelos termodinâmicos robustos e estratégias confiáveis de estimativa de parâmetros. A mistura binária ciclohexano–etanol é um exemplo de tal sistema, com grande importância para a indústria química do etanol, onde sua separação permanece como um passo crucial na produção de etanol anidro. Neste estudo, parâmetros de contribuição de grupo para o modelo UNIFAC foram otimizados utilizando o algoritmo de Nelder–Mead (método do simplex descendente), implementado no MATLAB, com base em dados de VLE isotérmicos da literatura (formalismo γ–φ) na faixa de temperatura de 278–338 K. Em seguida, os equilíbrios de fase isobáricos (40–150 kPa) foram avaliados e comparados com dados experimentais, com atenção especial à diferença de desempenho em relação ao modelo UNIFAC-Do (Dortmund), cujos parâmetros foram obtidos do banco de dados DDB. A influência do chute inicial no processo de otimização também foi investigada, assim como o efeito de distribuições alternativas de dados de coeficiente de atividade obtidos a partir de cálculos COSMO-SAC (concentrados na região rica em ciclohexano, na região rica em etanol ou uniformemente distribuídos ao longo da faixa de composição). Por fim, uma estratégia híbrida de otimização, combinando ciclos sucessivos de otimização, foi proposta e testada, revelando-se a mais equilibrada entre as abordagens investigadas, uma vez que superou os parâmetros ajustados isotermicamente sob condições isobáricas e apresentou desempenho superior ao modelo UNIFAC-Do na maioria dos casos isotérmicos, ao mesmo tempo em que alcançou um desempenho

complementar na reprodução dos coeficientes de atividade. Durante todos os cálculos, a equação de estado de Peng–Robinson, com regras clássicas de mistura de van der Waals, foi empregada para calcular os coeficientes de fugacidade na fase vapor. A comparação entre os parâmetros UNIFAC otimizados e os resultados do UNIFAC-Do destacou o compromisso entre previsões isotérmicas e isobáricas, bem como a capacidade da rotina híbrida de melhorar a consistência geral. Quanto à previsão do azeótropo, a metodologia proposta possibilitou uma reprodução satisfatória da temperatura e da composição a 101,3 kPa. Vale ressaltar que o formalismo adotado aqui não havia sido previamente aplicado ao sistema ciclohexano–etanol, reforçando assim a novidade e relevância desta contribuição.

Palavras-chave

Termodinâmica computacional; Peng- Robinson; UNIFAC; Pressão de bolha.

Table of Contents

List of Figures

List of Tables

1. Introduction

Vapor–liquid equilibrium (VLE) data and trustworthy predictive models of phase behavior constitute foundational elements for the design, optimization, and safe operation of separation processes across the chemical, petrochemical, and pharmaceutical industries [1]. Distillation remains one of the most ubiquitous and economically significant unit operations for effecting vapor–liquid separations; its rigorous design and reliable operation rely explicitly upon accurate thermodynamic descriptions of phase equilibria to size equipment, estimate energy consumption, determine feasible separation sequences, and guarantee product quality and process safety [2-4]. In the case of binary and multicomponent mixtures that deviate from Raoult's-law ideals, precise thermodynamic modelling is not a mere convenience but an operational necessity, since non-ideal behavior, frequently arising from specific intermolecular interactions such as hydrogen bonding, stark polarity contrasts, or marked differences in molecular size and shape, can produce azeotropes or complex, temperature- and pressure-sensitive phase boundaries that severely constrain the feasibility of conventional fractionation techniques [5-7].

Azeotropic behavior exemplifies the practical and conceptual challenges associated with strong non-idealities: an azeotrope is a composition at which the vapor and liquid phases become compositionally identical, making traditional fractional distillation unable to effect complete component separation [8]. From an engineering perspective, the presence of azeotropes compels the use of specialized separation sequences, azeotropic or extractive distillation, the deployment of selective entrainers, membrane or pervaporation processes, or the adoption of hybrid technological solutions [9, 10]. A salient industrial instance of such tactics is the dehydration of ethanol: molecular sieves and azeotropic/entrainer-based distillation remain economically important approaches, and entrainers such as cyclohexane are historically and practically employed to modify azeotropic loci and thereby enable the efficient production of anhydrous ethanol [11, 12]. The cyclohexane–ethanol binary system thus occupies an important niche where

academic interest and industrial relevance intersect: it represents a prototypical alkane–alcohol mixture exhibiting strong non-ideal behavior, a temperature-dependent azeotrope, and mechanistic features that illuminate the influence of size and polarity mismatch on phase equilibria.

Given the technological and economic importance of non-ideal and azeotropic systems, the availability of high-quality experimental VLE data and robust thermodynamic models is indispensable. When sufficient and well-distributed experimental data are accessible across the pertinent temperature and pressure ranges, regression of model parameters typically yields highly accurate predictions. Nevertheless, comprehensive experimental coverage is often incomplete: many binary pairs or specific thermodynamic windows remain sparsely characterized or unevenly sampled in composition–temperature–pressure space. This reality motivates two complementary strategies in contemporary thermodynamic practice: (i) the use of physically grounded, parametrized models that can be calibrated from limited data yet remain reliable across broader conditions [13], and (ii) the integration of theoretically computed thermodynamic quantities (e.g., activity coefficients produced by quantum-chemistry based continuum approaches) to augment or regularize parameter estimation when experimental information is lacking. Over the last several decades a broad palette of thermodynamic models has been developed to address these circumstances, ranging from cubic equations of state (CEoS) to local-composition activity-coefficient models and hybrid CEoS–GE mixing rules intended to combine vapor-phase fidelity with detailed liquid-phase non-ideality while keeping computational complexity manageable [14-16].

Among cubic equations of state, the Peng–Robinson (PR) formulation has achieved widespread adoption because of its concise algebraic structure, ease of implementation in process simulation environments, and reasonable performance for many systems, including moderately asymmetric mixtures and systems at elevated pressures [17, 18]. Nonetheless, the classical PR formulation exhibits recognized limitations when confronted with fluids showing strong hydrogen bonding, intense polarity, or severe molecular-size disparities [14, 18]. These limitations have motivated extensive refinements of the PR framework, including alternative cohesion α-functions, empirical modifications to the temperature

dependence of the attractive parameter a(T), and the adoption of advanced mixing rules that permit coupling the CEoS to excess-Gibbs (GE) models [19-25]. Such enhancements seek to preserve the numerical advantages of cubic EoS while extending their descriptive power for strongly non-ideal fluids.

Complementary to CEoS approaches, group-contribution activity-coefficient models such as UNIFAC and its Dortmund parameterization (UNIFAC-Do) continue to play a central role in the representation of liquid-phase non-ideality [26]. The UNIFAC methodology partitions molecules into functional groups, building the excess Gibbs energy from combinatorial and residual contributions derived from group surface areas, volumes and binary group interaction parameters [13, 27]; this group-centric philosophy provides two pragmatic advantages: (i) the ability to estimate activity coefficients for mixtures lacking comprehensive binary experimental data (predictive application), and (ii) the capacity to work with a reduced and transferable parameter set that can be refined by focused parameter estimation campaigns to reach higher fidelity. The Dortmund variant UNIFAC-Do extends the classical formulation through additional temperature-dependent terms and a larger suite of parameters intended to improve accuracy for strongly non-ideal mixtures, albeit at the cost of greater parameter-fitting effort and calibration care [28]. The γ–φ coupling, employing a cubic EoS for vapor-phase fugacities and a GE model for liquid-phase excess Gibbs energies, remains a pragmatic and widely used framework because it balances physical representativity and computational expedience [29].

Despite the existence of parameter databases such as the Dortmund Data Bank (DDB) [30, 31], default parameter sets are not universally optimal for every binary pair or thermodynamic window. Consequently, system-specific parameter estimation, fitting UNIFAC group interaction parameters to the most reliable experimental VLE data available, remains a well-established and necessary practice whenever high predictive fidelity is sought. However, parameter estimation can be impaired by sparse or unevenly distributed data; in response, computationally derived and thermodynamically consistent activity coefficients, obtained from approaches such as COSMO-SAC [32], have emerged as promising surrogates or supplements. COSMO-SAC (a statistical thermodynamics model founded on quantum-chemical σ-profiles) enables the estimation of activity coefficients across

broad composition and temperature ranges without requiring exhaustive experimental campaigns [33]. Modern, open implementations of COSMO-SAC and accessible sigma-profile databases permit researchers to compute activity coefficients at arbitrary compositions and temperatures [34, 35]; these computed activities can then be used to inform, augment or regularize the regression of group interaction parameters for UNIFAC-type models.

In the present dissertation, a pragmatic and integrative modeling philosophy is adopted. Pure-component vapor pressures and saturation behavior are described using the Peng–Robinson equation of state equipped with a fitted cohesion α-function, while liquid-phase non-idealities are modelled via UNIFAC-type frameworks (both the original UNIFAC and the Dortmund parametrization where appropriate). Parameter estimation for UNIFAC group interaction energies is pursued using several complementary strategies: (i) direct fitting to experimental isothermal VLE bubble-point data, (ii) fitting to activity-coefficient vectors generated by COSMO-SAC (thus exploiting theoretically-predicted, thermodynamically consistent data), and (iii) hybrid procedures that combine experimental VLE with COSMO-derived activities. These parallel approaches enable a systematic evaluation of (a) the extent to which COSMO-SAC data can meaningfully augment sparse experimental information, (b) the sensitivity of obtained parameter sets to the choice of initial guesses used in numerical optimization, and (c) the influence of composition sampling strategies (i.e., how the COSMO-generated activity-coefficient vector is constructed) and the temperature span included in the estimation procedure.

The cyclohexane–ethanol binary is especially well suited to such an investigation. It possesses direct industrial relevance (notably for ethanol dehydration and entrainer-based azeotropic distillation), exhibits prototypical alkane–alcohol non-idealities and an azeotrope whose location varies with temperature, and benefits from several experimental isothermal and isobaric VLE data sets published by different groups. This provides a useful basis for calibration and validation of model predictions; where experimental coverage is limited, thermodynamic activity data computed via COSMO-SAC (with an appropriate choice of sigma-profile database) are introduced into tailored optimization protocols to estimate UNIFAC group interaction parameters. The predictive

capability of the resulting parameter sets is evaluated quantitatively (average absolute deviations in bubble pressure, bubble temperature and vapor-phase composition) and graphically (*P*–*z* and *T*–*z* phase diagrams, together with azeotrope location).

The main contributions of this dissertation are as follows: (1) a comprehensive and up-to-date compilation and utilization of all available experimental isothermal and isobaric VLE data for the cyclohexane–ethanol binary; (2) the estimation and validation of Peng–Robinson cohesion α-function parameters for pure cyclohexane and ethanol using reference saturation data; (3) systematic parameter-estimation exercises for UNIFAC (with an assessment of UNIFAC-Do) using exclusively experimental VLE, exclusively COSMO-SAC activity coefficients, and hybrid strategies combining both data sources; (4) a detailed sensitivity analysis of the optimized parameters with respect to optimization initial guesses and to the composition sampling strategy of COSMO-SAC-generated activity vectors; and (5) a quantitative benchmarking of model performance against DDB/Dortmund parameterizations and the assembled experimental dataset. To the best of the author's knowledge, and subject to the exhaustive literature review presented in the subsequent chapter, the combined application of a Peng–Robinson EoS (with the selected α-function) and custom-fitted UNIFAC parameters for the cyclohexane–ethanol system, particularly when the UNIFAC parameters are informed by both experimental VLE and COSMO-SAC activity coefficients, has not been exhaustively examined in a single, unified study.

Finally, the dissertation is organized as follows. Chapter 2 defines the overarching research goals, motivating the scientific and industrial relevance of the work. Chapter 3 specifies the concrete objectives pursued, linking them directly to the methodological framework and expected contributions. Chapter 4 reviews the literature, with emphasis on thermodynamic modeling approaches, azeotropic distillation, and prior studies on alkane–alcohol systems, highlighting gaps that justify the present investigation. Chapter 5 summarizes the theoretical background and the computational tools and databases used. Chapter 6 details the methodology implemented for pure-component parameter estimation, computational VLE routines, and the optimization strategies (treatment of initial guesses, objective-function definitions and numerical routines). Chapter 7 contains the results and

discussion: pure-component saturation-pressure fits, isothermal and isobaric VLE modeling with both DDB and optimized UNIFAC parameter sets, COSMO-SAC-informed optimizations and sensitivity analyses, azeotrope predictions, and comprehensive comparisons. Chapter 8 presents the main conclusions, discusses limitations and outlines promising directions for future work.

2. Research Goals

The main goal of this dissertation is to estimate UNIFAC interaction parameters for the cyclohexane–ethanol system through bubble-point calculations using both VLE and COSMO-SAC activity coefficient data, and to identify the set of parameters that best represents the physical interactions in the system.

3. Specific Goals

- To estimate the Peng–Robinson α-function parameters for both pure components using NIST REFPROP saturation pressure data.
- To estimate UNIFAC parameters using only isothermal VLE data with different initial guesses.
- To estimate UNIFAC parameters using only COSMO-SAC activity coefficient data with different initial guesses.
- To estimate UNIFAC parameters using both VLE (isothermal and isobaric) and COSMO-SAC data with different initial guesses.

4. Literature Review

4.1. Vapor–liquid equilibrium studies on alkane–alcohol systems

Recent investigations on vapor–liquid equilibrium (VLE) of alkane–alcohol mixtures have provided valuable insights into the complex non-ideal behavior that characterizes these systems [36, 37]. In particular, detailed studies on n-octane combined with branched pentanol isomers have expanded the available experimental data for long-chain hydrocarbons interacting with medium-sized alcohols. Measurements were performed for four binaries (n-octane with 2-methyl-1-butanol, 2-methyl-2-butanol, 3-methyl-1-butanol, and 3-methyl-2-butanol) under atmospheric pressure, using a dynamic recirculating Gillespie-type still and gas chromatography for phase composition analysis. All mixtures exhibited temperature-minimum azeotropes and strong positive deviations from Raoult's law, consistent with the presence of significant excess Gibbs energies. Furthermore, the azeotropic composition was observed to shift toward the center of the compositional domain as the boiling-point difference between the constituents decreased, a trend aligned with earlier general observations for n-alkane/alcohol equilibria [36].

For the correlation and prediction of these data sets, several thermodynamic models were assessed. The Non-Random Two-Liquid (NRTL) model, with parameters regressed specifically for the studied binaries, delivered excellent correlations, with deviations between calculated and experimental values remaining very small across the entire composition range. Predictive group-contribution approaches were also tested. While the original UNIFAC formulation failed to reproduce the experimental phase behavior with acceptable accuracy, the modified UNIFAC (Dortmund) provided significant improvements. This superior performance is attributed to its inclusion of temperature-dependent interaction terms and its refined treatment of hydroxyl groups, which distinguishes primary, secondary, and tertiary alcohols. Nevertheless, even the Dortmund parameterization showed limitations for the n-octane/3-methyl-2-butanol system, where structural

features such as the presence of two methine groups appear to introduce specific interactions not fully captured by the existing group-interaction matrix.

The systematic comparison of NRTL, UNIFAC, and UNIFAC-Dortmund in this context highlights two key aspects. First, local-composition models, when supported by high-quality experimental data, remain the most reliable tools for accurate representation of a given binary system. Second, group-contribution models, though less accurate in a purely predictive mode, retain a critical advantage in transferability, especially when supported by extended parameter sets such as those in the Dortmund modification. However, the noted deficiencies for certain branched alcohols emphasize the importance of continuously refining group definitions and binary interaction parameters in order to enhance predictive reliability.

Complementary results have been obtained for another representative alkane–alcohol system, ethanol + n-heptane, studied under both isothermal and isobaric conditions [37]. Experimental measurements were conducted at three isotherms (363.15, 393.15, and 423.15 K) and at 101.33 kPa using a modified Rose–Williams equilibrium still coupled with an ebulliometer. The presence of an azeotrope was confirmed, and the data sets were subjected to thermodynamic consistency testing. For the correlation of the measured equilibria, the Peng–Robinson–Stryjek–Vera (PRSV) equation of state was employed, and different mixing rules were assessed. The Adachi–Sugie mixing rule proved to be the most suitable, producing an average absolute relative deviation (AARD) of 0.282% in bubble-point temperatures under isobaric conditions and 1.711% in isothermal VLE calculations. These results confirm that cubic equations of state, when combined with appropriately selected mixing rules, can provide accurate descriptions of alkane–alcohol phase equilibria across a wide temperature range.

In addition to VLE correlations, the ethanol–heptane study incorporated a thermophysical property analysis. The PRSV equation of state was combined with Eyring’s theory to correlate literature viscosity data for the mixture, achieving an average deviation of 2.81%. Although viscosity is not the primary focus of phase equilibrium studies, this result demonstrates the capacity of equation-of-state frameworks to extend predictive consistency beyond phase equilibria,

encompassing transport properties relevant to process design and biofuel applications such as injector and spray modeling [37].

Together, these works reinforce several important conclusions for the broader class of alkane–alcohol binaries. The formation of minimum-temperature azeotropes is recurrent and strongly linked to disparities in component volatilities. Local-composition models, when parameterized with system-specific data, provide excellent descriptive performance, while predictive group-contribution approaches require careful consideration of parameter sets and may demand additional refinements for branched or structurally complex alcohols. On the other hand, cubic equations of state, particularly PRSV with suitable mixing rules, remain an attractive option for systems investigated at elevated temperatures or for cases where a compact, transferable model framework is desired.

From a methodological perspective, both studies illustrate best practices in the characterization of non-ideal liquid mixtures: rigorous experimental validation, consistency testing, comparison of multiple modelling frameworks, and reporting of quantitative error metrics. The combined insights emphasize that the adequacy of a chosen thermodynamic model is highly context-dependent, varying with molecular structure, temperature range, and intended application. For alkane–alcohol mixtures, accurate representation of hydrogen bonding and dispersive asymmetry remains the critical challenge that dictates whether local-composition models, group-contribution methods, or cubic equations of state will provide the most reliable predictions.

4.2. Review of previous experimental and modelling studies on the cyclohexane–ethanol binary

A considerable body of experimental and modelling work has been devoted to the cyclohexane–ethanol binary mixture, motivated primarily by its technological importance as an entrainer system for ethanol dehydration and, more generally, by its status as a prototypical alkane–alcohol pair that exhibits pronounced non-ideal behaviour and an azeotrope whose location is sensitive to temperature and pressure. The literature that directly addresses this specific binary is useful not only because it provides reference data for model calibration, but also because it exposes the strengths and limitations of different thermodynamic

descriptions when applied to alkane–alcohol systems. The following paragraphs synthesise the principal efforts that have measured phase equilibria for this pair and/or applied common activity-coefficient and equation-of-state models to represent its behaviour [38-45].

Early experimental and modelling studies combined classical excess-Gibbs approaches with virial or empirical vapour-phase treatments. For instance, Scatchard and Satkiewicz [17] reported extensive measurements and employed a hybrid modelling strategy (Wilson for the liquid phase combined with a virial description for the vapour) across a range of temperatures. Their work demonstrated good agreement between calculation and experiment for the majority of the composition range, average deviations in pressure were reported to be very small (below 0.1% for most compositions). These results nonetheless highlighted a recurring practical limitation: in the extreme dilute limits (i.e., when one component is present at trace concentrations) deviations increase, reaching values on the order of a few percent (≈2% in the dilute extremes reported by Scatchard and Satkiewicz). Such sensitivity in dilute regions is typical for mixtures where molecular asymmetry and specific interactions (hydrogen bonding from the alcohol) exert a large effect on activity coefficients, and it underscores the importance of dense experimental sampling near the ends of the composition scale when high-accuracy predictions are required [38].

More recent comparative modelling studies have systematically assessed the performance of a broad suite of liquid-phase activity models applied to the cyclohexane–ethanol system. Joseph, Ramjugernath and Raal [41] carried out a comprehensive evaluation using a battery of excess-Gibbs formalisms (Wilson, modified Wilson, 4-suffix Margules, van Laar, NRTL, UNIQUAC and modified UNIQUAC) enabling a direct, model-to-model comparison against the same experimental dataset. Their findings indicate that relatively simple, two-parameter models such as Wilson can, in certain temperature ranges, reproduce the observed VLE with remarkable fidelity; Joseph *et al*. report an average absolute deviation in pressure not exceeding 0.25 kPa for the Wilson formulation in the conditions studied. Such outcomes illustrate that parsimonious models, when carefully parameterised, can sometimes match or outperform more elaborate representations for specific systems or data ranges [41].

Complementary studies at atmospheric pressure (101.3 kPa) have produced slightly differing rankings of model performance, demonstrating that model adequacy is often context-dependent. For example, in the work attributed to Shujie Li and co-workers [39], the authors compared NRTL, UNIQUAC and Wilson, assuming ideal gas behaviour in the vapour phase, and reported that NRTL produced marginally lower root-mean-square deviations (RMSD = 0.0160) than Wilson (RMSD = 0.0240), whereas UNIQUAC performed substantially worse (RMSD = 0.1262) for the specific dataset and fitting-procedure they used. This type of result highlights two important points: (i) objective metrics (RMSD, AAD, etc.) depend sensitively on the subset of experimental points used in regression and on the weighting scheme adopted in the objective function, and (ii) different parameter estimation strategies and underlying assumptions (e.g., ideal vapour assumption versus explicit equation-of-state treatment) can materially affect the apparent ranking of model performance [39].

Other authors have adopted classical polynomial or empirical excess-property correlations as practical, data-driven ways to fit experimental VLE. Keshpande *et al*. (often cited as Koshpande or Keshpande *et al*. in the literature) [40] fitted the Redlich–Kister expansion for the excess molar Gibbs energy to observed VLE and reported good qualitative agreement with measurements at 101.3 kPa. While polynomial approaches such as Redlich–Kister are powerful for interpolation and can provide compact parameter sets, they are inherently descriptive rather than predictive; their transferability outside the fitted thermodynamic window is limited compared with physically motivated models such as NRTL or UNIFAC. Keshpande *et al*. did not report an extensive quantitative error analysis in the particular publication referenced here, a fact that limits a rigorous, apples-to-apples comparison with studies that do report AAD or RMSD metrics [40].

Taken together, these studies illustrate several recurring themes that are directly relevant to the present dissertation. First, a variety of liquid-phase models, from empirical polynomial correlations to local-composition models and group-contribution schemes, can be tuned to reproduce cyclohexane–ethanol VLE with differing degrees of success, and model rankings are not universal but depend on the dataset, the objective function used in regression, and the thermodynamic assumptions adopted for the vapour phase. Second, dilute limits and extreme

composition regions remain challenging for many models, producing comparatively larger deviations; this is a practical concern for entrainer design, where small compositional changes can have outsized effects on azeotrope location and separation performance. Third, because different studies employ different subsets of available experimental data (and sometimes different experimental protocols), there is a tangible benefit in assembling a comprehensive, up-to-date compilation of all trustworthy isothermal and isobaric measurements for a given binary, a task that this dissertation explicitly undertakes and uses as the benchmark against which alternative parameter-estimation strategies are tested [38-41].

Finally, reviewing the available literature indicates an opportunity and a gap that motivate the methodology developed in this work. While the classical activity models and empirical correlations discussed above have been widely applied to cyclohexane–ethanol, to the best of authors' knowledge, there are no prior publications that explicitly combine a cubic equation of state (for the vapour phase) with a UNIFAC parameterization refined for the cyclohexane–ethanol binary. Moreover, and equally importantly, we are not aware of any study that has employed COSMO-SAC-derived activity-coefficient vectors as a primary source of thermodynamically consistent data to augment sparse experimental measurements during the estimation of UNIFAC group interaction parameters for this particular system. This absence in the literature motivates a systematic and integrative investigation that: (i) fits and validates Peng–Robinson pure-component cohesion (α-function) parameters against reference saturation data, (ii) derives UNIFAC group interaction parameters using alternative strategies (direct regression to experimental isothermal/isobaric VLE, calibration against COSMO-SAC activity vectors, and hybrid combinations of both data types) and (iii) interrogates the sensitivity of the resulting parameter sets to initial guesses, to the sampling strategy adopted for COSMO-SAC composition vectors, and to the temperature range included in the optimizations. Such a structured programme of work aims to produce parameter sets that are both physically grounded and practically predictive for an industrially relevant, strongly non-ideal system such as cyclohexane–ethanol, thereby addressing an explicit void in the existing body of knowledge. [38-41].

4.3. Use of COSMO-SAC activity-coefficient data in VLE modelling: the monochlorobenzene – 4,6-dichloropyrimidine case study

The study by Haaz *et al*. [46] presents the inaugural isobaric vapor–liquid equilibrium measurements for the binary system monochlorobenzene (MCB) + 4,6-dichloropyrimidine (DCP) and deliberately couples these laboratory results with first-principles, density-functional COSMO-SAC predictions to appraise the feasibility of a wholly predictive route when empirical data are scarce. Framed from a process-engineering vantage, DCP being an industrially relevant heterocycle and MCB its manufacture solvent, the work positions high-quality experimental VLE as complementary to quantum-chemistry-based forecasting, the joint aim being to supply reliable phase-equilibrium inputs for the design of purification and solvent-recovery operations.

Experimentation was executed on a modified Gillespie isobaric still with meticulous attention to purity, sampling and analytics: both components underwent vacuum distillation, compositions were obtained by GC-FID (with refractometric validation where applicable), and the apparatus incorporated duplicated cooled sampler walls, Cotrell circulation and vigorous stirring to ensure phase homogeneity. Reported measurement uncertainties (±0.1 K in temperature, ±2 kPa in pressure) together with a Herrington area consistency test ($\Delta H - JH = 5.3\%$) substantiate the internal thermodynamic consistency and overall quality of the dataset obtained at 101 kPa across the full composition range.

On the computational side the authors generated single-conformer geometries and performed density-functional optimizations to produce σ-profiles and segment definitions that feed the COSMO-SAC segment-based activity-coefficient formalism; for benchmarking they also evaluated the conventional UNIFAC approach. This pragmatic combination of microscopic σ-profile descriptors and macroscopic VLE computation allows a direct assessment of how well a quantum-chemistry rooted model can reproduce engineering-relevant phase behaviour without empirical parameterization.

Quantitatively, COSMO-SAC reproduces the measured T–x–y loci with striking fidelity: the authors report an overall deviation on the order of 3.0×10^{-3} and objective-function metrics of 9.9×10^{-4} for temperature, 6.3×10^{-3} for liquid

mole fraction x_1 and 1.7×10^{-3} for vapor mole fraction y_1, with residuals increasing as $x_1 \rightarrow 0$ (the very DCP-rich/dilute regime). By contrast, UNIFAC, while capturing the general y(x) trend, exhibits noticeably larger discrepancies in the intermediate composition window (approximately $0.05 < x_1 < 0.4$) and a degraded capacity to follow the experimental temperature curve T(x,y), thereby attenuating its utility for precise thermal-design calculations in this non-ideal, halogenated/heterocyclic system.

From a process-design perspective the combined evidence leads to two practical inferences: no azeotrope is present in the examined range, so conventional distillative separation of DCP from MCB remains viable, and the high degree of agreement between COSMO-SAC and experiment supports the pragmatic deployment of first-principles calculations to fill gaps in physicochemical databases during early-stage design. The authors, prudently, temper this endorsement with the caveat that model refinement may be necessary and predictive performance can be chemical-family dependent, but the present case study nonetheless furnishes compelling, engineering-grade support for integrating COSMO-SAC into workflows where experimental data are limited.

5. Theoretical Foundations

5.1. Thermodynamic Model – Peng–Robinson Equation of State

The prediction of vapor–liquid equilibrium (VLE) properties in this work relied initially on the Peng–Robinson (PR) cubic equation of state, a widely employed model in process simulation and phase behavior studies due to its balance between computational simplicity and reasonable accuracy for a broad range of compounds. In particular, the PR model is especially suited for nonpolar and slightly polar systems, while also serving as a robust reference framework when coupled with excess Gibbs energy models for nonideal mixtures.

The general expression of the Peng–Robinson equation of state can be written as:

$$P = \frac{RT}{V_m - b} - \frac{a\alpha}{V_m^2 + 2bV_m - b^2} \tag{1}$$

where P is the system pressure, T is the absolute temperature, R is the universal gas constant, V_m is the molar volume, and a, b, and α are the parameters that account for intermolecular interactions and the non-ideality of the fluid.

The attractive parameter a and the co-volume parameter b are defined as follows:

$$a = a_C \alpha \tag{2}$$

$$b = 0.0778 \frac{RT_C}{P_C} \tag{3}$$

$$a_C = 0.45724 \frac{R^2 T_C^2}{P_C} \tag{4}$$

where T_C and P_C represent the critical temperature and critical pressure of the pure component, respectively.

The temperature-dependent factor α, which corrects the attractive term to account for the deviation of real fluids from ideal behavior, was calculated using a generalized expression proposed by Almeida, Aznar, and Telles (1991) [22]:

$$\alpha = exp\left[\lambda_1(1 - T_r)|1 - T_r|^{(\lambda_2 - 1)} + \lambda_3\left(\frac{1}{T_r} - 1\right)\right] \quad (5)$$

where the reduced temperature is defined as:

$$T_r = \frac{T}{T_C} \quad (6)$$

The coefficient λ_i (i = 1, 2, 3) are component-specific parameters, which in this work were initially taken from literature sources (André Young, Fernando L. P. Pessoa, *et al.*, 2017) [47]. These parameters govern the curvature of the $\alpha(T_r)$ function, thereby allowing a better representation of the temperature dependence of the attractive forces for each substance.

Through this formulation, the Peng–Robinson model provides a solid theoretical foundation for calculating bubble pressures and generating baseline thermodynamic predictions, which are subsequently refined by integration with activity coefficient models for highly nonideal mixtures.

5.2. The *γ-φ* Approach for Vapor–Liquid Equilibrium Calculations

Isothermal phase diagrams for the binary cyclohexane–ethanol system were obtained by calculating bubble and dew-point curves over a range of liquid compositions. The construction of these diagrams relies on the fundamental thermodynamic criterion for phase equilibrium, which requires that the fugacity of component *i* in the vapor phase equals its fugacity in the liquid phase, as expressed in Equation (7).

$$f_i^v = f_i^l \quad (7)$$

Within the γ-φ framework, however, the liquid-phase fugacity is conveniently expressed as the product of the activity coefficient and mole fraction, while the vapor-phase fugacity is obtained through an equation of state. Accordingly, the equilibrium condition can be rewritten as Equation (8):

$$\phi_i^v y_i P = \gamma_i^l x_i f_i^{sat} \quad (8)$$

Here, ϕ_i^v denotes the fugacity coefficient of component *i* in the vapor phase, y_i the corresponding vapor-phase mole fraction, P the total system pressure, γ_i^l the activity coefficient of component *i* in the liquid phase, x_i the liquid-phase mole fraction, and f_i^{sat} the fugacity of pure component *i* at saturation. It should be

emphasized that, in the present study, the Poynting correction factor was assumed equal to unity, as the operating pressures were insufficiently high for its contribution to exert a significant influence.

The evaluation of fugacity coefficients for the vapor phase requires the application of mixing rules to determine the equation-of-state parameters *a* and *b* for mixtures. Among the various mixing rules reported in the literature, the classical van der Waals mixing rules were adopted in this work, as defined in Equations (9-11):

$$a_{mix} = \sum_i \sum_j x_i x_j a_{ij} \tag{9}$$

$$b_{mix} = \sum_i x_i b_i \tag{10}$$

$$a_{ij} = \sqrt{a_{ii} a_{jj}}\left(1 - k_{ij}\right) \tag{11}$$

In these expressions, k_{ij} is an empirical binary interaction parameter. For the cyclohexane–ethanol system, k_{ij} was set to zero, since preliminary sensitivity analyses revealed that its effect on the phase behavior predictions was negligible.

To implement the γ-φ methodology, a dedicated MATLAB code was developed, employing a bubble-pressure algorithm inspired by the procedure outlined by Sandler [48]. The program iteratively computes equilibrium by varying the bubble-point compositions and applying the equilibrium condition defined in Equation (8). Pressure is subsequently updated based on the deviation between the computed sum of vapor-phase mole fractions and its theoretical value of unity. Within this framework, the fugacity coefficients of vapor-phase species were evaluated using the Peng–Robinson equation of state, whereas the liquid-phase activity coefficients were calculated using the UNIFAC and UNIFAC-Dortmund models.

5.3. Group Interaction Parameters for UNIFAC and UNIFAC-Do Models

The UNIFAC model comprises a combinatorial part ($\ln \gamma_i{}^C$), which represents the contribution of molecular shape and size, and a residual part ($\ln \gamma_i{}^R$), which describes the contribution of binary group interaction energy (Eq. 12).

$$\ln \gamma_i = \ln {\gamma_i}^C + \ln {\gamma_i}^R \quad (12)$$

The application of this model requires knowledge of specific parameters. The group volume (R_k) and surface area (Q_k) parameters of each component are essential for calculating the combinatorial contribution to the activity coefficient. The full formulation of this combinatorial part can be found in Fredenslund, Jones, and Prausnitz [13]. The R_k and Q_k values for each group were obtained from the DDB database [31] and are listed in the Appendix (Table A.1).

Furthermore, group interaction parameters are used to calculate the residual part of the activity coefficient. For the cyclohexane–ethanol system, six interaction parameters are required when using the UNIFAC model, while eighteen parameters are needed for UNIFAC-Do. These group interaction parameters were also obtained from the DDB database [31] and are presented in Table A.2 (for UNIFAC) and Table A.3 (for UNIFAC-Do). The formulation of the residual part is shown in equations 13 to 16 [13], where equation 15 applies to the UNIFAC model, and equation 16 applies to the UNIFAC-Do model. As shown, UNIFAC-Do includes both linear and quadratic temperature-dependent terms.

$$\ln {\gamma_i}^R = \sum_k {\nu_k}^{(i)} \left[\ln \Gamma_k - \ln {\Gamma_k}^{(i)}\right] \quad (13)$$

Here, ${\nu_k}^{(i)}$ is the number of groups of type k in component i, Γ_k is the group residual activity coefficient, and ${\Gamma_k}^{(i)}$ is the residual activity coefficient of group k in a solution containing only component i.

$$\ln \Gamma_k = Q_k \left[1 - \ln\left(\sum_m \Theta_m \Psi_{mk}\right) - \sum_m \left(\frac{\Theta_m \Psi_{km}}{\sum_n \Theta_n \Psi_{nm}}\right)\right] \quad (14)$$

$$\Psi_{mn} = exp\left(-\frac{a_{mn}}{T}\right) \quad (15)$$

$$\Psi_{mn} = exp\left(\frac{-(a_{mn} + b_{mn}T + c_{mn}T^2)}{T}\right) \quad (16)$$

In these equations, Θ_m represents the area fraction of group m, Ψ is the group interaction factor, and a_{mn}, b_{mn} and c_{mn} are adjustable group interaction parameters.

6. Methodology

6.1. Optimization of the Adjustable Parameters of the Cohesion α Function

To enable the calculation of vapor pressures for pure cyclohexane and ethanol, a computational routine was developed in MATLAB based on the Peng–Robinson equation of state. This widely adopted equation requires the optimization of certain adjustable parameters for each component. The number of adjustable parameters depends on the chosen cohesion α function.

In the present work, the form of the α function proposed by Almeida, Aznar, and Telles [22] was adopted. This function introduces three empirical parameters (denoted by λ); its explicit mathematical form is presented in Section 5.1.

The numerical strategy used to determine the vapor pressures involves iteratively correcting the temperature estimate by comparing fugacity coefficients in the liquid and vapor phases. This iterative approach was inspired by the algorithmic procedures described by Sandler [48], which are known for their robustness in vapor–liquid equilibrium calculations.

The adjustable parameters λ_1, λ_2 and λ_3 in the adopted α function were estimated by minimizing the deviation between calculated and experimental vapor pressures for pure cyclohexane and ethanol. For this purpose, the Nelder–Mead optimization algorithm (also known as the downhill simplex method), as described in Lagarias *et al.* [49, 50], was employed, as implemented in MATLAB's built-in function *fminsearch*. Experimental data for vapor pressure were obtained from the Dortmund Data Bank (DDB) [30], while the critical properties T_C and P_C used in the calculations were taken from the National Institute of Standards and Technology (NIST) database [51].

Initial estimates for the parameters of the α function were based on values reported by André Young, Fernando L. P. Pessoa, *et al*. [47] for both compounds.

The objective function minimized during the parameter estimation process was defined as:

$$F = \sum_{i} \left| \frac{P_{EXP,i} - P_{CALC,i}}{P_{EXP,i}} \right| \quad (17)$$

In this equation, $P_{EXP,i}$ is the experimental vapor pressure at the i-th temperature, and $P_{CALC,i}$ is the calculated vapor pressure at the same temperature. After optimization, the estimated parameters were used to compute the saturation pressure of each component. The results were compared with experimental vapor pressure data from the DDB [30] as well as with data obtained using REFPROP (version 10.0), developed by NIST as a reference database for fluid thermodynamic and transport properties [52].

6.2. Optimization of Group Interaction Parameters for the UNIFAC Model

As discussed in the previous section, the accurate modeling of vapor pressure for pure components is essential for calculating phase equilibria involving mixtures. Building upon the reliable estimation of parameters used in the calculation of saturation pressures for cyclohexane and ethanol, the next step in this study focuses on the application of activity coefficient models to represent non-ideal behavior in binary mixtures. Among the most widely adopted models for this purpose is UNIFAC, which estimates activity coefficients based on molecular group contributions. In contrast to the UNIFAC-Do variant, which uses a more complex temperature-dependent formulation and a larger number of parameters, the original UNIFAC model offers a simpler alternative with fewer adjustable parameters, making it attractive for the development of customized parameter sets tailored to specific binary systems.

While default parameters are available from the Dortmund Data Bank (DDB) for both the original UNIFAC and the extended UNIFAC-Do models, these values are not necessarily optimized for the cyclohexane–ethanol system. In this context, a parameter estimation procedure was conducted to optimize the group interaction parameters of the UNIFAC model specifically for the cyclohexane–ethanol binary system. There were two main motivations for this effort: (i) to reduce the total

number of group interaction parameters, and (ii) to achieve accurate representation of the vapor–liquid equilibrium (VLE) behavior for the system under study.

The optimization was performed using experimental isothermal VLE data available in the literature [38, 41]. As initial guesses for the optimization, the standard group interaction parameters extracted from the Dortmund Data Bank (DDB) were used. These values are presented in Table 1 below:

Table 1 - Group interaction parameters for the UNIFAC model in the cyclohexane - ethanol system, extracted from DDB database (optimization initial guesses)

i	j	a_{ij}	a_{ji}
CH_3/CH_2	OH	986.5	156.4
CH_3/CH_2	Cy-CH_2	-450.4	-34.36
OH	Cy-CH_2	-817.7	1913

The structural information required to calculate the combinatorial contribution of the UNIFAC model—namely, the group volume (R) and surface area (Q) parameters—was also taken from the DDB and kept constant throughout the optimization procedure. These values are shown in Table 2:

Table 2 - Group volume (R) and surface area (Q) parameters for each molecular group in the cyclohexane–ethanol system, extracted from DDB database

Molecular Groups	R	Q
CH_3	0.6325	1.0608
CH_2	0.6325	0.7081
OH	1.2302	0.8927
Cyclic-CH_2	0.7136	0.8635

The group interaction parameters for the UNIFAC-Do model, which were used later for comparative analysis, are presented in Table 3:

Table 3 - Group interaction parameters for the UNIFAC-Do model in the cyclohexane–ethanol system extracted from DDB database.

i	j	a_{ij}	b_{ij}	c_{ij}	a_{ji}	b_{ji}	c_{ji}
CH_3/CH_2	OH	2777	-4.674	0.001551	1606	-4.746	0.0009181
CH_3/CH_2	Cy-CH_2	-117.1	0.5481	-0.00098	170.9	-0.8062	0.001291
OH	Cy-CH_2	3121	-13.69	0.01446	2601	-1.25	-0.006309

The optimization of the UNIFAC group interaction parameters was performed using MATLAB's *fminsearch* function, which implements the Nelder–Mead simplex algorithm. The objective function used in the optimization, shown in Equation (18), accounts for the relative deviations in both bubble pressure and vapor-phase composition across multiple temperatures and mixture compositions:

$$F = \sum_T \sum_i \left(\left| \frac{P_{EXP,i,T} - P_{CALC,i,T}}{P_{CALC,i,T}} \right| - \left| \frac{y_{EXP,i,T} - y_{CALC,i,T}}{y_{EXP,i,T}} \right| \right) \quad (18)$$

In this expression, $P_{EXP,i,T}$ and $P_{CALC,i,T}$ are the experimental and calculated bubble pressures, respectively, for the i-th global composition at temperature T. Similarly, $y_{EXP,i,T}$ and $y_{CALC,i,T}$ represent the experimental and calculated vapor-phase mole fraction of cyclohexane under the same conditions.

In order to simplify the optimization process and reduce the number of parameters being simultaneously estimated, a key assumption was adopted: both CH_3 and CH_2 groups were treated as belonging to the same interaction group "family". Consequently, identical energy interaction values were used for these groups, regardless of the group with which they were interacting. This simplification is consistent with the parameter structure employed in the DDB database, as reflected in Table 1.

Upon completion of the optimization, the new set of estimated parameters was used to generate isothermal phase diagrams for the cyclohexane–ethanol system. These diagrams were then evaluated through visual comparison with experimental data, and quantitatively assessed using the average absolute deviations in both pressure and vapor-phase composition. The predictive performance of the optimized UNIFAC model was compared to that of the UNIFAC-Do model, providing insight into the trade-offs between model simplicity and accuracy, as will be explored in the subsequent sections.

6.3. Modeling of Isobaric Phase Diagrams for the Cyclohexane–Ethanol Binary System

In order to evaluate the modeling accuracy of the group interaction parameters (previously estimated based on isothermal data) under conditions not considered during their optimization, isobaric phase diagrams were then calculated

and compared with those obtained using Table 3 (UNIFAC-Do) values, also employing absolute average deviation in bubble temperature and vapor-phase composition as comparison metrics. In this way, the accuracy of the parameter set estimated in this study was benchmarked against the DDB database.

Experimental isobaric VLE data for the cyclohexane–ethanol binary system were collected at five different pressures: 40 kPa, 70 kPa, 98 kPa, 101.3 kPa, and 150 kPa. The data at 40 kPa, 70 kPa, 98 kPa, and 150 kPa were extracted from the study by Reddy *et al.* [43], while data at atmospheric pressure (101.3 kPa) were gathered from three different sources: Zhao *et al.* [44], Keshpande *et al*. [40], and Li *et al* [42].

For constructing these isobaric phase diagrams, a MATLAB program was developed to calculate the bubble temperature of the mixture, using a code similar to that previously employed for bubble pressure calculations, as suggested by Sandler [48]. The code performs calculations by varying the bubble-point compositions using the equilibrium criterion of the γ–φ method (Eq. 8), followed by updating the temperature based on the difference between the sum of the vapor-phase mole fractions and its expected value of 1.

6.4. Development of the MATLAB Routine for Optimization Using COSMO-SAC

In order to improve the reliability and accuracy of the obtained parameter sets, additional approaches were investigated for the optimization of UNIFAC group interaction parameters. Among these, the COSMO-SAC model was employed to generate activity coefficient data, which were subsequently incorporated into the parameter estimation procedure.

Through a collaborative effort with researchers from the Technical University of Berlin (Technische Universität Berlin – TU Berlin), it was initially possible to obtain activity coefficient data for each component of the cyclohexane–ethanol binary system at three distinct temperatures: 278 K, 308 K, and 338 K. Notably, these temperatures are part of the set of six temperatures previously used throughout the parameter optimization studies conducted in this work.

For the optimization of the UNIFAC group interaction parameters (as described in Eq. 15) within this newly developed MATLAB routine, the following objective function was employed to minimize the deviation between the activity coefficients predicted by the UNIFAC model and those obtained from the COSMO-SAC model:

$$F = \sum_{T} \sum_{i} \frac{\left| \gamma_{1,CALC,i,T} - \gamma_{1,COSMO,i,T} \right|}{\gamma_{1,COSMO,i,T}} \tag{19}$$

In this equation, $\gamma_{1,CALC,i,T}$ represents the activity coefficient of component 1 (cyclohexane) calculated according to the complete methodology described in Section 5.3 of the present work, whereas $\gamma_{1,COSMO,i,T}$ corresponds to the activity coefficients of cyclohexane obtained using the COSMO-SAC model through the collaboration with the Berlin research group. Both coefficients refer to the same cyclohexane mole fraction *i* and temperature T. The activity coefficients for ethanol were not included in the objective function because, according to the Gibbs–Duhem relation, the activity coefficients of the two components in a binary mixture are not independent. Since the COSMO-SAC data are thermodynamically consistent, the variation in the activity coefficient of one component inherently determines the variation of the other, making it redundant to include both in the optimization procedure.

6.5. Analysis of the Effect of Initial Guess on the VLE and COSMO-SAC Optimization Routines

In any parameter optimization process, regardless of its complexity or the number of parameters involved, the ultimate goal is to obtain a set of optimized values that, ideally, is independent of the initial guess provided at the beginning of the procedure. With this objective in mind, the present work includes an analysis of how different initial guesses influence the outcome of the optimization process.

This analysis was carried out using two distinct MATLAB routines, with two different initial guesses tested for each. The first MATLAB routine examined was the one developed to model the vapor–liquid equilibrium (VLE) behavior of the cyclohexane–ethanol binary system by means of bubble pressure calculations, as previously described in Section 5.2 of this work. In this routine, the group

interaction parameters of the UNIFAC model were optimized (as formulated in Eq. 15), using an objective function designed to minimize the deviations in both bubble pressure and vapor-phase mole fraction, as detailed in Eq. 18. For this evaluation, two different initial guesses were employed to initiate the optimization procedure.

The initial guesses adopted in this first routine were as follows: (i) the original UNIFAC group interaction parameters provided by the DDBST database (which are listed in Table 1), and (ii) a manually perturbed version of the same DDBST parameters, in which each interaction value was randomly modified to introduce slight deviations. The modified values used as the second initial guess are shown below in Table 4.

Table 4 - Manual modification of DDBST parameters used as initial guess.

i	j	a_{ij}	a_{ji}
CH_3/CH_2	OH	686.5	356.4
CH_3/CH_2	Cy-CH_2	-150.4	-134.36
OH	Cy-CH_2	-517.7	913

The second MATLAB routine investigated in this analysis was the recently developed one that incorporates activity coefficient data generated by the COSMO-SAC model into the optimization procedure, as outlined in Section 6.4. In this routine, the UNIFAC group interaction parameters (again following the formulation given in Eq. 15) were optimized by minimizing the relative deviation between the activity coefficients of cyclohexane obtained from COSMO-SAC and those calculated using the methodology described in Section 5.3. The corresponding objective function used for this optimization is presented in Eq. 19. Similarly to the first routine, two distinct initial guesses were also tested for this second optimization procedure.

The initial guesses adopted in this second case consisted of (i) the original UNIFAC group interaction parameters available on the DDBST website (as presented in Table 1), and (ii) a set of interaction parameters previously estimated using isothermal experimental data collected at six different temperatures, as a result of the optimization procedure discussed in Section 6.2.

It is important to emphasize that, for both MATLAB routines analyzed in this section (the one modeling VLE through bubble pressure calculations and the one employing COSMO-SAC activity coefficient data) the optimization processes using the different initial guesses described above were conducted simultaneously at three different temperatures: 278 K, 308 K, and 338 K. These specific temperatures were chosen because they correspond to the same set of temperatures for which activity coefficient data for each component of the cyclohexane–ethanol binary system were obtained through collaboration with researchers from the Technical University of Berlin. It is also worth noting that all optimization procedures carried out throughout this work, including both routines described above, were performed using the native MATLAB function *fminsearch*, which implements the Nelder–Mead simplex algorithm, also known as the downhill simplex method.

6.6. Optimization Using COSMO-SAC Activity Coefficient Data at All Six Temperatures

Following the acquisition of the COSMO-SAC algorithm, which was developed in Python and made freely available on the GitHub platform through the article *"A Benchmark Open-Source Implementation of COSMO-SAC"* by Ian H. Bell, Jadran Vrabec *et al.* [32], it became possible to compute activity coefficient data for each component of the cyclohexane–ethanol binary mixture at all six temperatures previously adopted throughout this work (278 K, 293 K, 308 K, 313 K, 323 K, and 338 K). To perform these calculations, it is necessary to select a sigma-profile database, since a few available options are provided within the GitHub repository upon downloading the source code. In this study, the sigma-profile database developed by the University of Delaware was chosen, as it yielded superior results in preliminary optimization tests when compared to other available databases [53].

The optimization of the UNIFAC group interaction parameters was also carried out using the *fminsearch* MATLAB built-in function, which minimizes the relative deviation between the activity coefficients of cyclohexane calculated according to the methodology outlined in Section 5.3 and those obtained from the COSMO-SAC model, as described in the objective function presented in Equation 19. In this step, the optimization was conducted using data from the full set of six temperatures.

In this stage of the study, an additional analysis was conducted to investigate whether the specific structure of the activity coefficient vector obtained from the COSMO-SAC algorithm has any influence on the optimization outcomes. In other words, when using the COSMO-SAC implementation in Python, it is possible to compute activity coefficients for each component across a wide range of mole fractions of one of the species, in this case, the mole fraction of cyclohexane was selected as the reference variable. This process results in a vector of activity coefficient values that represent the thermodynamic behavior of the system at various compositions.

This vector of activity coefficients may be constructed in different ways: the data points can be evenly spaced across the entire composition range, or alternatively, they can be concentrated either toward the region corresponding to high cyclohexane mole fractions (i.e., the more concentrated region) or toward the more dilute region (i.e., low cyclohexane concentrations). These different approaches for generating the activity coefficient vectors are illustrated in Figure 1 below.

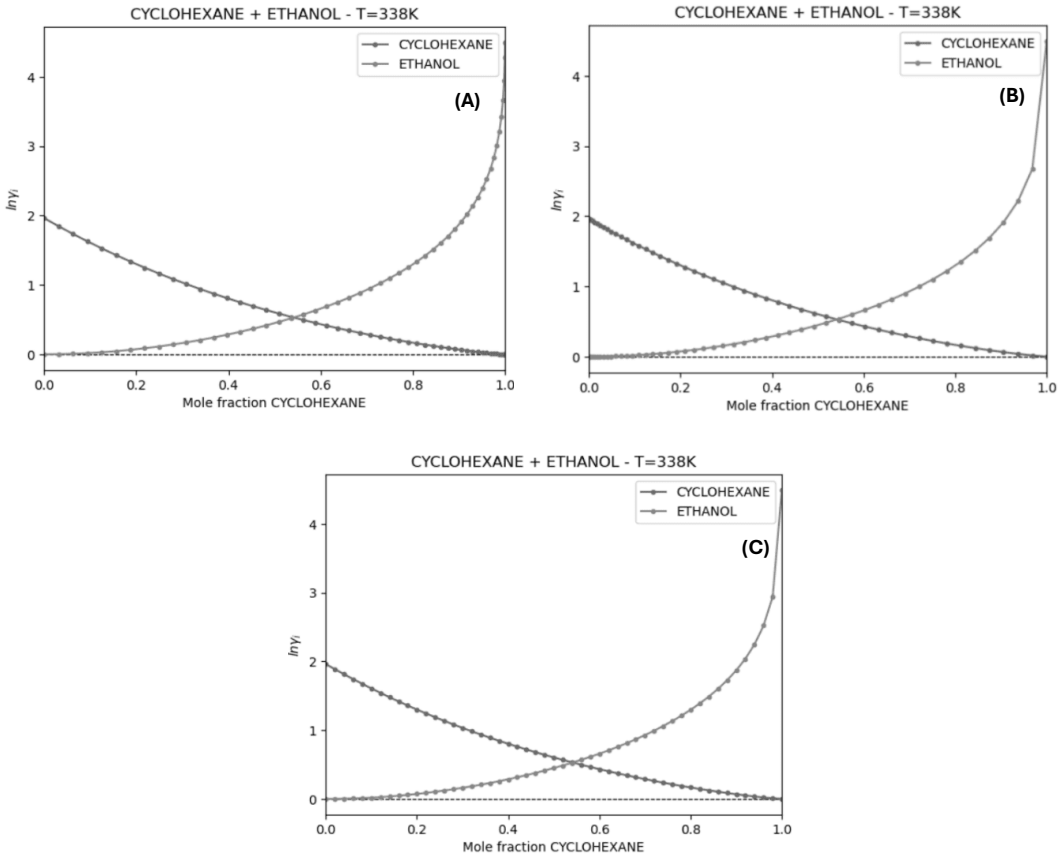

Figure 1 - Activity coefficient data obtained using the COSMO-SAC model for the cyclohexane–ethanol binary system at a temperature of 338 K. (A) The activity coefficient values are more densely distributed in the composition range corresponding to higher concentrations of cyclohexane. (B) The activity coefficient values are more densely distributed in the composition range corresponding to lower concentrations of cyclohexane (i.e., the dilute region). (C) The activity coefficient values are evenly spaced across the entire range of cyclohexane mole fractions.

After the parameter estimation procedure was completed, the UNIFAC group interaction parameters obtained for each of the different scenarios (namely, activity coefficient vectors concentrated toward the region of higher cyclohexane mole fractions, concentrated toward the more dilute region, and uniformly spaced across the entire composition range) were employed to calculate both isothermal and isobaric phase diagrams for the cyclohexane–ethanol binary system.

These resulting phase diagrams were then compared with one another in order to assess whether the aforementioned variations in the structure of the activity coefficient vectors used during optimization had any significant impact on the parameter estimation process or on the predictive quality of the model.

6.7. New Parameter Estimation Strategy: A Hybrid Approach

In order to enhance the parameter estimation process and to obtain a set of group interaction parameters capable of more accurately describing both isothermal and isobaric phase diagrams, as well as reliably predicting activity coefficients, a final hybrid strategy was employed in the parameter optimization procedure.

This proposed strategy consists of two successive optimization stages. In the first stage, a cycle of optimization of the UNIFAC model's group interaction parameters is carried out, following a similar general strategy previously described in Section 6.4. As in that section, the objective was to minimize the relative deviation between the calculated activity coefficients of cyclohexane and those obtained from the COSMO-SAC model. However, in this new strategy, a slightly modified objective function is adopted, one based on absolute deviations rather than relative deviations, as expressed in the equation below:

$$F = \sum_{T} \sum_{i} \left| \gamma_{1,CALC,i,T} - \gamma_{1,COSMO,i,T} \right| \tag{20}$$

In this equation, $\gamma_{1,CALC,i,T}$ represents the activity coefficient of cyclohexane calculated using the current UNIFAC parameters, whereas $\gamma_{1,COSMO,i,T}$ refers to the corresponding value predicted by the COSMO-SAC model, both evaluated at global mixture composition *i* and temperature T.

This modification was motivated by an important limitation observed in the previously employed objective function based on relative deviations. In the case of relative deviations, the calculated error associated with each data point is normalized by the value of the reference activity coefficient predicted by COSMO-SAC. However, in strongly non-ideal regions, particularly under conditions where cyclohexane is present in trace amounts, its activity coefficient becomes significantly greater than unity. Therefore, the denominator in the relative deviation term becomes large, causing the contribution of that deviation to the objective function to be artificially reduced. This effect undesirably weakens the weight of deviations precisely in the highly diluted regime, where the activity coefficients are most sensitive and deviations are potentially more meaningful. To address this drawback, the new strategy adopts a formulation based on absolute deviations,

which treats all deviations with equal weight regardless of the magnitude of the reference value.

Upon completion of this first optimization stage, the predictive capability of the estimated parameter set is quantitatively assessed in terms of its ability to reproduce experimental VLE data and activity coefficients for both components of the mixture. Only after this evaluation is the parameter set passed to the second stage of the optimization strategy, where it serves as the initial guess for a second cycle of refinement. In this subsequent stage, a new objective function is employed, which includes the deviations between COSMO-SAC and UNIFAC activity coefficients, as well as deviations between calculated and experimental bubble pressures, and vapor-phase compositions of cyclohexane corresponding to isothermal (*P–z*) phase diagrams. The complete objective function used in this second cycle is presented below:

$$F = \sum_{T}\sum_{i}\left|\gamma_{1,CALC,i,T} - \gamma_{1,COSMO,i,T}\right| + \left|P_{CALC,i,T} - P_{EXP,i,T}\right| + \left|y_{P,1,CALC,i,T} - y_{P,1,EXP,i,T}\right| \quad (21)$$

In this expression, $\gamma_{1,CALC,i,T}$ and $\gamma_{1,COSMO,i,T}$ are, respectively, the calculated and COSMO-SAC activity coefficients of cyclohexane; $P_{CALC,i,T}$ and $P_{EXP,i,T}$ represent the calculated and experimental bubble pressures; $y_{P,1,CALC,i,T}$, $y_{P,1,EXP,i,T}$ denote the calculated and experimental vapor-phase mole fractions of cyclohexane for the isothermal equilibrium; All these quantities are evaluated at the global mixture composition *i* and temperature T. It is worth emphasizing that, in both optimization cycles, only absolute deviations were considered in the objective functions, and that all optimization calculations described in this section were also performed using the *fminsearch* function available in MATLAB.

7. Results and Discussion

7.1. Saturation Pressure Calculation for Pure Cyclohexane and Ethanol

To assess the reliability of the estimated parameters for the pure substances involved in this study, we began by modeling the saturation pressures of cyclohexane and ethanol using the Peng–Robinson equation of state, enhanced by the Almeida–Aznar–Telles formulation for the α function. The specific form of the cohesion function employed in this context is defined in Eq. 5. The three adjustable parameters (λ_1, λ_2, λ_3) associated with this formulation were optimized for each pure component to ensure the best possible agreement with reference data.

These optimized parameter values are compiled in Table 5, which provides a concise overview of the λ values obtained for both cyclohexane and ethanol. They serve as the foundation for the saturation pressure predictions discussed below.

Table 5 - Estimated Adjustable Parameters (λ) for the Almeida–Aznar–Telles Cohesion α Function.

Component	λ_1	λ_2	λ_3
Cyclohexane	0.5986	1.0053	0.0799
Ethanol	1.6105	1.079	-0.1457

With these parameters in hand, the saturation pressure of each pure compound was calculated over a wide temperature range and compared with highly accurate reference values retrieved from the NIST REFPROP software. The results of these comparisons are graphically presented in Figures 2A and 2B. As illustrated, the predicted pressures closely follow the reference data throughout the temperature range, demonstrating strong quantitative agreement for both substances. This result reinforces the adequacy of the parameter set derived for the present modeling framework.

To provide a more quantitative measure of the model's accuracy, the average absolute deviation (AAD) in saturation pressure was calculated for each compound, according to the expression given in Eq. 22:

$$AAD = \frac{1}{N}\sum_{i=1}^{N}\left|P_{REFPROP,i} - P_{CALC,i}\right| \tag{22}$$

In this equation, $P_{REFPROP,i}$ is the saturation pressure obtained from the REFPROP software at temperature *i* and $P_{CALC,i}$ is the saturation pressure calculated at the same temperature. The resulting AAD values were found to be 0.091 bar for cyclohexane and 0.076 bar for ethanol, which further validates the high fidelity of the fitted parameters for the pure components of interest.

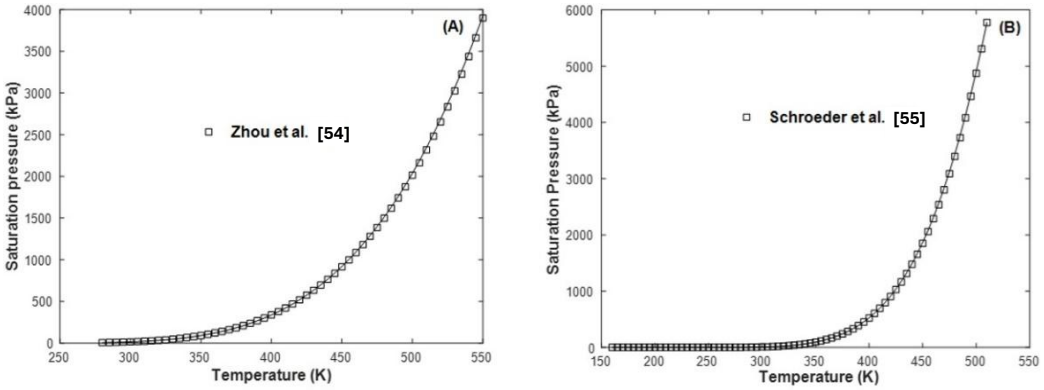

Figure 2 - Calculated saturation pressure of cyclohexane (A) and ethanol (B) as a function of temperature.

7.2. Modeling Isothermal VLE with DDB Parameters for UNIFAC-Do

In order to assess the predictive capabilities of the UNIFAC-Do model when employing group interaction parameters sourced from the DDB database [31], bubble pressure calculations were carried out to generate $P{-}z$ (pressure–global composition) diagrams at each of the six temperatures for which experimental data are available in the literature for the cyclohexane–ethanol system. These temperatures (278 K, 293 K, 308 K, 313 K, 323 K, and 338 K) cover a broad thermal range, thereby enabling a more comprehensive evaluation of the model's performance across different thermodynamic regimes.

The results of these isothermal vapor–liquid equilibrium (VLE) calculations are partially presented in Figure 3, which displays the $P-z$ phase diagrams for the two limiting temperatures: 278 K and 338 K. The diagrams corresponding to the intermediate temperatures (293 K, 308 K, 313 K, and 323 K) are available in the Appendix (Figures A.1A to A.1D). A visual comparison between the calculated and experimental bubble pressures and vapor-phase compositions reveals that the UNIFAC-Do model, when using the DDB parameters, provides a reasonable description of the phase behavior throughout the entire temperature range examined.

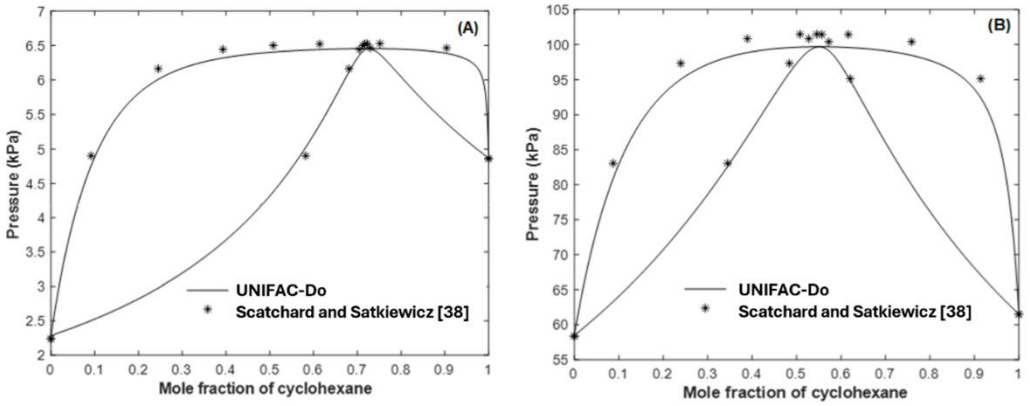

Figure 3 - $P-z$ phase diagrams at (A) 278 K and (B) 338 K for the cyclohexane - ethanol system using DDB parameters for UNIFAC-Do.

Nevertheless, it is possible to observe a trend: the agreement between the model predictions and the experimental data tends to be better at lower temperatures (278 K, 293 K, and 308 K), while slightly larger deviations are observed at higher temperatures (313 K, 323 K, and 338 K). This trend is further corroborated by the quantitative data presented in Table 6, which reports the average absolute deviations (AAD) in both bubble pressure and vapor-phase mole fraction of cyclohexane at each temperature. The equations used for these calculations are provided below for reference.

$$AAD_P = \frac{1}{N}\sum_{i=1}^{N}\left|P_{EXP,i} - P_{CALC,i}\right| \tag{23}$$

$$AAD_y = \frac{1}{N}\sum_{i=1}^{N}\left|y_{EXP,i} - y_{CALC,i}\right| \tag{24}$$

Where $y_{EXP,i}$ and $y_{CALC,i}$ denote the experimental and calculated vapor-phase mole fraction of cyclohexane for the *i*-th global composition value.

Table 6 - Average absolute deviations in bubble pressure (AAD_P) and vapor-phase composition of cyclohexane (AAD_y) for isothermal VLE calculations using DDB parameters with the UNIFAC-Do model.

Temperature (K)	AAD_P (kPa)	$AAD_y .10^{-2}$
278	0.11	1.01
293	0.25	0.86
308	0.48	0.89
313	1.21	0.93
323	1.07	0.84
338	2.11	1.01

Altogether, the graphical and numerical results suggest that the UNIFAC-Do model with DDB parameters is capable of reasonably reproducing isothermal VLE data for the cyclohexane–ethanol system, particularly under lower-temperature conditions. However, some discrepancies remain, especially at elevated temperatures, indicating potential limitations in the model's ability to capture non-ideal behavior in such cases. Nevertheless, improvements could still be achieved, particularly at higher temperatures, as will be explored in the subsequent optimization procedures.

7.3. Isothermal VLE Modeling with Optimized UNIFAC Parameters

The group interaction parameters estimated for the UNIFAC model in this work, specifically optimized for the cyclohexane–ethanol binary system, are presented below:

Table 7 - Estimated UNIFAC Group interaction parameters for the cyclohexane–ethanol system.

i	j	a_{ij}	a_{ji}
CH_3/CH_2	OH	5881.4	6923.4
CH_3/CH_2	Cy-CH_2	-13.3	75.1
OH	Cy-CH_2	344.2	1634.2

Using these optimized parameters, isothermal vapor–liquid equilibrium (VLE) calculations were carried out. The resulting pressure–composition (*P–z*) diagrams are shown in Figure 4, which includes the limiting cases of 278 K and 338

K. The remaining diagrams for intermediate temperatures (293 K, 308 K, 313 K, and 323 K) are provided in the Appendix (Figures A.2A to A.2D).

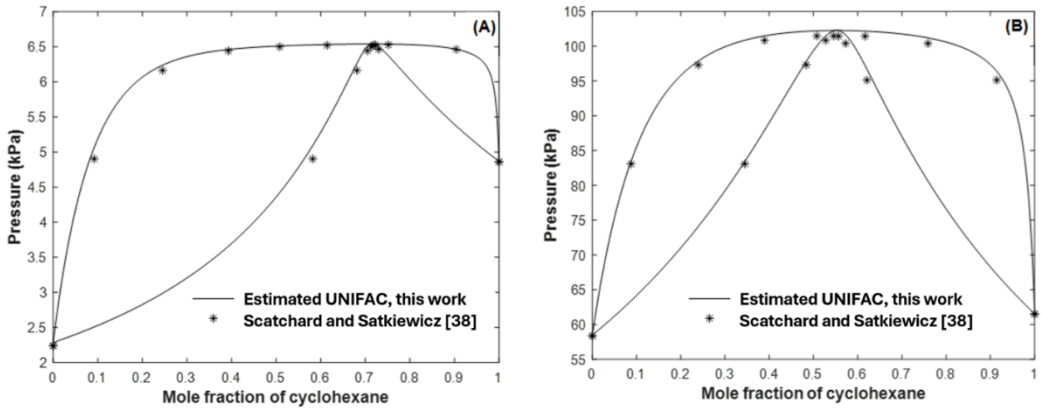

Figure 4 - *P*−*z* phase diagrams at (A) 278 K and (B) 338 K for the cyclohexane–ethanol system using the group interaction parameters estimated in this work for the UNIFAC model.

A visual inspection of the diagrams reveals that the new set of parameters significantly improves the agreement between calculated and experimental phase behavior, particularly in capturing the pressure variation as a function of composition across the entire temperature range. This improvement becomes even more evident when comparing these results (Figure 4) to those obtained using the UNIFAC-Do model with standard parameters from the Dortmund Data Bank (Figure 3). While the original UNIFAC-Do model exhibited noticeable deviations in certain composition and temperature ranges, the optimized UNIFAC model demonstrates a more consistent and accurate reproduction of the experimental data.

This qualitative assessment is further supported by the quantitative metrics presented in Tables 8 and 9, which list the average absolute deviations (*AAD*) in bubble pressure (*AADP*) and in the vapor-phase mole fraction of cyclohexane (*AADY*), respectively. In all temperature conditions examined, the *AAD* values associated with the optimized parameters are markedly lower than those of the UNIFAC-Do model, reaffirming the enhanced predictive capability achieved through the parameter estimation process developed in this work.

Table 8 - Average absolute deviations in bubble pressure (AAD_P) for isothermal VLE calculations using DDB parameters for the UNIFAC-Do model and the parameters estimated in this work for the UNIFAC model.

Temperature (K)	AAD_P (kPa)	
	UNIFAC-Do	This work
278	0.11	0.04
293	0.25	0.06
308	0.48	0.18
313	1.21	0.52
323	1.07	0.11
338	2.11	0.83

Table 9 - Average absolute deviations in cyclohexane vapor-phase composition (AAD_Y) for isothermal VLE calculations using DDB parameters for the UNIFAC-Do model and the parameters estimated in this work for the UNIFAC model.

Temperature (K)	$AAD_y .10^{-2}$	
	UNIFAC-Do	This work
278	1.01	0.56
293	0.86	0.30
308	0.89	0.83
313	0.93	0.21
323	0.84	0.18
338	1.01	0.20

It is worth noting that the standard parameters available in the Dortmund Data Bank for the UNIFAC-Do model (shown in Table A.3) are not necessarily specific to the cyclohexane–ethanol binary system. Instead, they are likely derived from correlations involving a large number of systems sharing similar functional groups. This generalized approach, although useful in many contexts, may not be adequate for providing the most accurate description of a particular system's phase behavior, especially when high fidelity is desired, as in the present work. This may partly explain the inferior performance of the UNIFAC-Do model when compared to the customized parameter set proposed here.

Regarding the broader context of VLE modeling for this binary system, it is instructive to consider the results obtained by Joseph, Ramjugernath, and Raal [41], who conducted one of the most detailed investigations on both isothermal and

isobaric VLE data for cyclohexane and ethanol. Their comprehensive study included experimental measurements at multiple conditions and evaluated the performance of several activity coefficient models, including Wilson, modified Wilson, 4-suffix Margules, van Laar, NRTL, UNIQUAC, and modified UNIQUAC. For each model, the authors reported average deviations between experimental and calculated pressures and vapor-phase compositions.

Among the various models examined, the Wilson model provided the best agreement with experimental data, yielding *AADP* = 0.09 kPa and *AADY* = 0.007 at 313 K, and *AADP* = 0.25 kPa and *AADY* = 0.006 at 323 K. When compared with the values reported in Tables 8 and 9, it becomes evident that the Wilson model slightly outperforms the present UNIFAC model at 313 K. However, at 323 K, the optimized UNIFAC parameters proposed here result in a lower *AADP* and *AADY*, suggesting a more accurate description at this higher temperature.

It is also important to point out that in their study, Joseph, Ramjugernath, and Raal estimated Wilson parameters based directly on the experimental data collected at 313 and 323 K. By contrast, the parameter estimation procedure developed in this work relied on experimental data from a broader temperature range, including 278, 293, 308, and 338 K. This approach not only enhances the robustness of the estimated parameters but also contributes to their applicability across a wider range of operating conditions. Nevertheless, improvements could still be achieved, especially at the higher temperatures, as will be explored in the following sections through parameter refinement and optimization strategies.

7.4. Modeling of Isobaric Vapor–Liquid Equilibrium at Multiple Pressures

In order to evaluate the capability of both parameter sets (namely, the UNIFAC-Do parameters obtained from the Dortmund Data Bank (DDB) and the set estimated in this work) for representing isobaric vapor–liquid equilibrium (VLE), a series of *T*–*z* diagrams were generated under five different pressure conditions and subsequently compared to available experimental data.

The modeling results corresponding to pressures of 40 kPa and 150 kPa are presented in Figures 5 and 6, respectively. For the remaining pressures investigated (70 kPa, 98 kPa, and 101.3 kPa), the corresponding diagrams are provided in the

Appendix (Figures A.3A to A.3F), so as to maintain a clear and organized presentation of the main text.

Visual inspection of Figures 5 and 6 reveals that both parameterizations are capable of capturing the qualitative features of the isobaric phase behavior of the system. However, a superior agreement with the experimental data was observed for the UNIFAC-Do model employing DDB parameters, particularly in terms of accurately reproducing the curvature of the bubble lines. Nonetheless, the optimized parameter set proposed in this work also demonstrated a reasonably good performance, considering its significantly lower complexity.

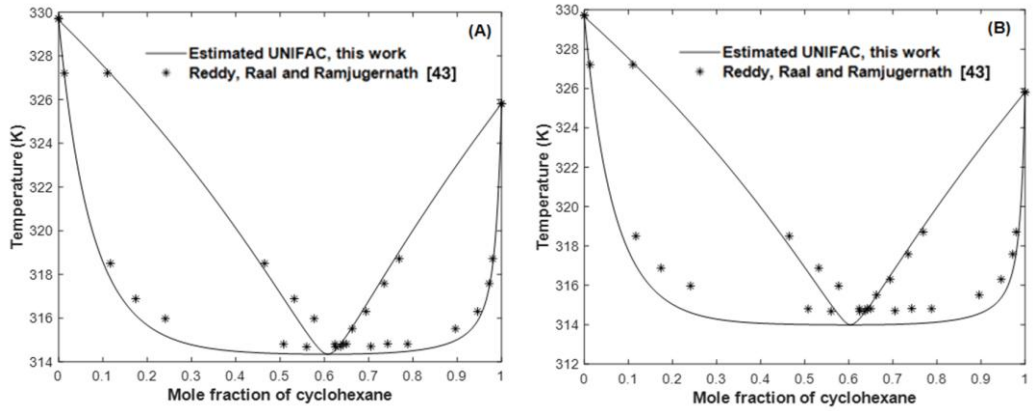

Figure 5 - Isobaric $T-z$ diagrams calculated using DDB UNIFAC-Do parameters (A) and optimized UNIFAC (B) at 40 kPa.

To quantitatively assess the performance of both models, average absolute deviations in bubble temperature (AAD_T) and in vapor-phase composition of cyclohexane (AAD_Y) were computed for each pressure condition. The AAD_T is defined by the following equation:

$$AAD_T = \frac{1}{N}\sum_{i=1}^{N}\left|T_{EXP,i} - T_{CALC,i}\right| \tag{25}$$

where $T_{EXP,i}$ and $T_{CALC,i}$ represent the experimental and calculated bubble temperature at global composition i, respectively.

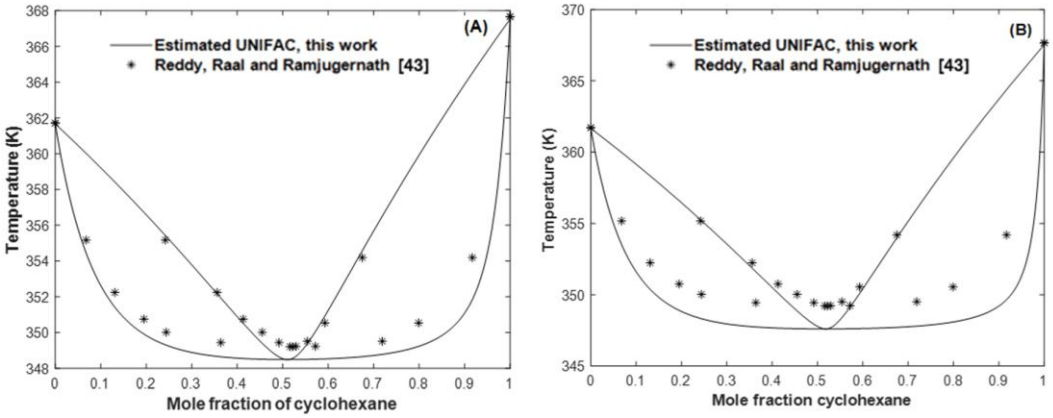

Figure 6 - Isobaric T−z diagrams calculated using DDB UNIFAC-Do parameters (A) and optimized UNIFAC (B) at 150 kPa.

The AAD_T and AAD_Y values obtained for each pressure are summarized in Tables 10 and 11. The results indicate that the eighteen-parameter version of the UNIFAC-Do model exhibits superior predictive accuracy under isobaric conditions across all pressures analyzed.

Table 10 - Average absolute deviations in bubble temperature (AAD_T) for isobaric VLE modeling using DDB parameters for the UNIFAC-Do model and the parameters estimated in this work for the UNIFAC model.

Pressure (kPa)	AAD_T (K)	
	UNIFAC-Do	This work
40	0.43	0.93
70	0.69	1.48
98	0.72	1.73
101.32	0.65	1.48
150	0.98	2.07

Table 11 - Average absolute deviations in cyclohexane vapor-phase composition (AAD_y) for isobaric VLE modeling using DDB parameters for the UNIFAC-Do model and the parameters estimated in this work for the UNIFAC model.

Pressure (kPa)	$AAD_y .10^{-2}$	
	UNIFAC-Do	This work
40	1.79	2.70
70	1.81	3.17
98	1.82	3.27
101.32	1.80	2.81
150	2.36	3.49

Present results indicate that, although the optimized parameter set developed in this work performs better for isothermal diagrams, the UNIFAC-Do model using DDB parameters shows a better performance in modeling isobaric diagrams. The superior performance from DDB dataset at isobaric conditions and lower performance at isothermal conditions suggest that the construction of DDB dataset regarding the groups covered in present was based mainly on isobaric experimental data assessment and not isothermal. In present work, the opposite was made, and isothermal data were used as a basis for estimating the parameters involved, and, indeed, a much better performance in describing the VLE behavior of the binary was achieved. On the other hand, at isobaric conditions, present proposed dataset, although enabled a good description of VLE behavior, especially regarding bubble temperature calculations, higher deviations from the available experimental data were observed in comparison with DDB eighteen parameter dataset.

Even so, it is important to emphasize that the optimized parameter set developed in this study is markedly simpler, comprising only six parameters, and does not incorporate any empirical temperature-dependence terms, which are included in the original UNIFAC-Do formulation. Thus, despite its simplicity, the present parameterization still achieved a fairly good representation of the isobaric VLE behavior, particularly for bubble temperature predictions, highlighting its potential as a practical alternative in cases where simplicity and computational efficiency are prioritized.

7.4.1. Azeotropic point calculation at 101.3 kPa

Since azeotropic mixtures pose a challenge for conventional distillation, accurate understanding and modeling of their azeotropic behavior holds significant technological relevance for the full separation of mixture components. Regarding the cyclohexane-ethanol system, azeotropic distillation plays a particularly relevant role in applications where ethanol separation is essential, such as in the production of anhydrous ethanol [11]. In this context, cyclohexane has been employed in azeotropic distillation processes, shifting the azeotropic point in T - z space and enabling the complete separation of components [12]. To evaluate the accuracy of the azeotropic point prediction for this system, the present study compares experimental data with calculations obtained from the UNIFAC and UNIFAC-Do

models at atmospheric pressure (101.3 kPa). The results of this comparative analysis are summarized in Table 12.

Table 12 - Comparison of experimental and calculated azeotropic points for the cyclohexane-ethanol system at atmospheric pressure.

Reference	Azeotrope Temperature (K)	Azeotrope composition
Present work (UNIFAC)	336.86	0.544
UNIFAC-Do	337.53	0.542
Experimental [19,21,23]	337.92	0.566

Both the UNIFAC and UNIFAC-Do models demonstrated satisfactory performance in predicting the azeotropic point of the cyclohexane-ethanol system at 101.3 kPa. However, the UNIFAC-Do model yielded a result slightly closer to the experimental azeotrope temperature, with a relative deviation of 0.12%, compared to 0.31% obtained with the UNIFAC model using the parameters estimated in the present study. Conversely, regarding the azeotrope composition, the UNIFAC model with the estimated parameters performed marginally better, exhibiting a relative deviation of 3.82%, in contrast to 4.32% observed for the UNIFAC-Do model. Unfortunately, a similar comparison could not be performed at other pressures, as the limited number of equilibrium data around the azeotrope did not allow for an accurate estimation of its location.

7.5. Sensitivity of UNIFAC Optimization to Initial Guesses: Comparative Assessment of the VLE and COSMO-SAC Routines

7.5.1. VLE-Based Optimization Routine

7.5.1.1. Initial Guess: Original UNIFAC Parameters from the DDBST Database

To evaluate the impact of the initial guess on the parameter optimization process, the MATLAB routine developed to simulate the vapor–liquid equilibrium (VLE) of the cyclohexane–ethanol binary mixture, via bubble pressure calculations as described in Section 6.2, was employed using two different starting points for the UNIFAC group interaction parameters. The first initial guess consisted of the original UNIFAC parameters available in the DDBST database, presented in Table

1 (Section 6.2). The results obtained after performing successive optimizations using this initial guess are compiled in Table 13 below.

Each row in Table 13 represents a new optimization cycle, where the parameter set obtained from the previous optimization was used as the starting point for the subsequent one. For every parameter set, the corresponding value of the objective function (formulated in Equation 18 to minimize deviations in bubble pressure and vapor-phase composition) is also reported. As previously stated in Section 2.5, this analysis was restricted to three temperatures: 278 K, 308 K, and 338 K.

Table 13 - Estimated UNIFAC group interaction parameters after successive optimization cycles using the original DDBST parameter set as initial guess. Optimizations were performed using MATLAB's *fminsearch* function (Nelder–Mead simplex algorithm).

Parameters set	**Estimated UNIFAC Group Interaction Parameters**						**Objective Function**
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	986.5	-450.4	156.4	-817.7	-34.36	1913	28.84
1st Optimization	6543.90	226.70	627.32	-8.75	-11.62	1639.27	0.390
2nd Optimization	3118.45	204.29	1563.44	32.53	-2.17	1668.54	0.248
3rd Optimization	2443.12	200.55	1868.71	7.95	12.32	1648.69	0.244
4th Optimization	18146.61	212.93	1981.43	-14.43	18.09	1630.22	0.238
5th Optimization	20439.08	210.76	2235.83	-13.73	18.91	1629.83	0.237
6th Optimization	3620560.53	207.99	1837.53	-16.36	20.98	1653.03	0.237
7th Optimization	7222299.75	208.65	2079.83	-17.59	21.91	1633.25	0.236

In the notation used, subscripts 1, 2, and 3 correspond to the functional groups present in the system: "1" refers to the methyl/methylene group (CH_3/CH_2), "2" denotes the hydroxyl group (OH), and "3" represents the cyclic CH_2 group of cyclohexane.

The results show that, from the second optimization cycle onward, markedly different parameter sets yield very similar values for the objective function. For example, the objective function value for the third optimization is only 0.006 higher than that obtained in the fourth cycle, despite the interaction parameters being drastically different. This suggests the existence of multiple local minima in the parameter space, with shallow objective function landscapes where distinct parameter sets produce nearly equivalent fits to the experimental VLE data.

Moreover, a clear lack of convergence is observed in the parameter values. As successive optimizations proceed, the interaction parameters tend to diverge and assume increasingly extreme values, even though the improvement in the objective function becomes marginal. This behavior indicates that the algorithm is attempting to exploit flat regions of the optimization surface, where significant changes in parameter values yield only negligible gains in model performance. Such instability raises concerns about the identifiability and physical relevance of the optimized parameter sets.

For these reasons, the optimization process was halted after the seventh cycle, as further reductions in the objective function became insignificant, while the resulting parameters continued to escalate beyond physically meaningful ranges.

7.5.1.2. Optimization Using Manually Perturbed DDBST Parameters as Initial Guess

The second initial guess employed in the parameter estimation procedure consisted of a manually adjusted version of the original DDBST parameter set, as presented in Table 4 in Section 6.5. The rationale behind this approach was to investigate how deliberate perturbations applied to a physically reasonable initial estimate would influence the optimization path and the resulting convergence behavior.

The outcomes of successive optimization cycles performed using this manually modified initial guess are summarized in Table 14. These optimizations were carried out using the *fminsearch* function from MATLAB, which is based on the Nelder–Mead simplex algorithm.

Table 14 - Estimated UNIFAC group interaction parameters after successive optimization cycles using a manually modified version of the DDBST parameters as the initial guess. All optimizations were performed using the Nelder–Mead simplex algorithm.

Parameters set	Estimated UNIFAC Group Interaction Parameters						Objective Function
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	686.5	-150.4	356.4	-517.7	-134.36	913	24.97
1st Optimization	2099.73	14.26	-79.16	-92.82	61.48	1500.24	0.432
2nd Optimization	1512.70	-2.43	-371.76	-115.29	-52.37	1449.91	0.288
3rd Optimization	2464.28	27.59	-448.32	-77.75	-165.61	1519.12	0.193
4th Optimization	160142.11	23.32	-436.59	-65.22	-162.23	1520.00	0.192
5th Optimization	169270.71	23.45	-436.48	-65.20	-162.33	1521.18	0.192
6th Optimization	236020.89	23.45	-436.49	-65.20	-162.33	1521.16	0.192
7th Optimization	76263.70	23.46	-436.49	-65.20	-162.33	1521.16	0.192

As observed in Table 14, the objective function experiences a pronounced decrease during the initial optimization steps. However, starting from the fourth optimization cycle, the reduction in the objective function value becomes negligible, despite substantial variations in the estimated parameters. This behavior indicates the presence of multiple parameter sets yielding comparable modeling performance in terms of the objective function, suggesting the existence of local minima or a relatively flat region in the objective function surface.

A particularly noteworthy observation emerges when comparing the sixth and seventh optimization results. Both parameter sets return identical objective function values (confirmed even at extended decimal precision), yet the parameter a_{12} exhibits a drastic difference between the two solutions. This highlights the potential for parameter redundancy or compensation effects, where variations in certain parameters are offset by adjustments in others, preserving the overall fit quality.

Additionally, parameters a_{13}, a_{21}, a_{23}, a_{31} and a_{32} appear to converge toward stable values throughout the optimization cycles. In contrast, parameter a_{12} fluctuates significantly, particularly in later stages, which may indicate a lower sensitivity of the objective function to this specific interaction parameter or a higher degree of freedom associated with its variation.

Given the stagnation in the objective function and the emerging redundancy in the solutions, the optimization process was interrupted after the seventh cycle.

Further iterations were deemed unnecessary, as they were unlikely to yield meaningful improvements or provide additional insights into the parameter estimation process.

7.5.1.3. Comparative Assessment of Optimized Parameters Obtained from Distinct Initial Guesses

In order to quantitatively evaluate the predictive accuracy of the final group interaction parameters estimated from the two distinct initial guesses (namely, the original DDBST parameters and their manually modified counterparts) a comparative analysis was conducted based on two performance metrics: the average absolute deviation in bubble pressure (AAD_P) and the average absolute deviation in the vapor-phase mole fraction of cyclohexane (AAD_Y). These quantities were calculated at each isothermal condition employed in the optimization procedure, and the results are summarized in Tables 15 and 16. The mathematical definitions of AAD_P and AAD_Y are provided in Equations (20) and (21), respectively.

Table 15 - Average absolute deviations in bubble pressure (AAD_P) obtained using final parameters estimated with different initial guesses in the VLE-based optimization routine.

Temperature (K)	AAD_P (kPa)	
	DDBST parameters as Initial Guess	*Manually modified DDBST Parameters as Initial Guess*
278	0.038	0.011
293	0.079	0.074
308	0.200	0.240
313	0.599	0.737
323	0.053	0.152
338	0.406	0.398

Table 16 - Average absolute deviations in the vapor-phase mole fraction of cyclohexane (AAD_Y) obtained using final parameters estimated with different initial guesses in the VLE-based optimization routine.

Temperature (K)	$AAD_y .10^{-2}$	
	DDBST parameters as Initial Guess	*Manually modified DDBST Parameters as Initial Guess*
278	0.629	0.513
293	0.392	0.367
308	0.841	0.887
313	0.266	0.389
323	0.273	0.217
338	0.308	0.153

It is important to emphasize that the parameters used to generate these results correspond to the final values obtained after seven consecutive optimization cycles in both cases, as presented previously in Tables 13 and 14. That is, each data point reflects the performance of the fully optimized model and not of the initial guess itself.

Subsequently, using these two distinct sets of estimated group interaction parameters, vapor–liquid equilibrium (VLE) calculations were carried out at all experimental temperatures. The corresponding pressure–composition (*P–z*) diagrams are presented in Figures 7 and 8 for the limiting temperatures of 278 K and 338 K, respectively. The results at intermediate temperatures (293 K, 308 K, 313 K, and 323 K) are provided in the Appendix (Figures A.4A–A.4H).

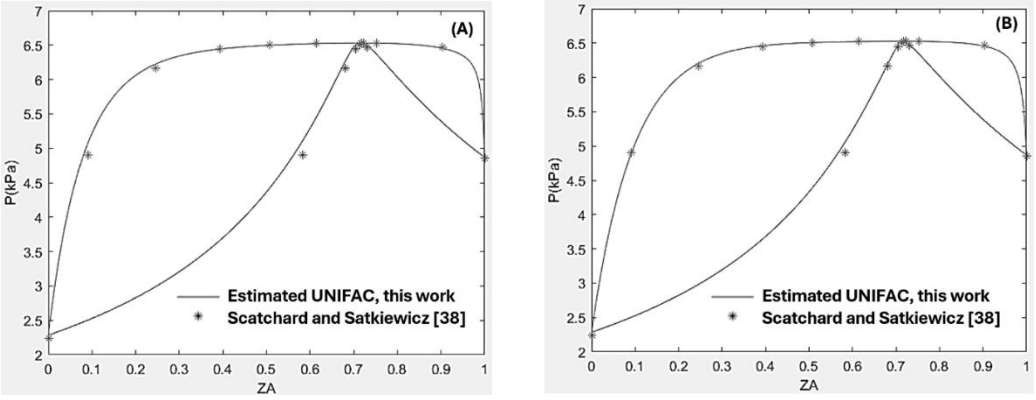

Figure 7 - Pressure–composition (*P–z*) diagram at 278 K comparing results obtained with group interaction parameters estimated using two different initial guesses: (A) the original DDBST values and (B) the manually modified DDBST values.

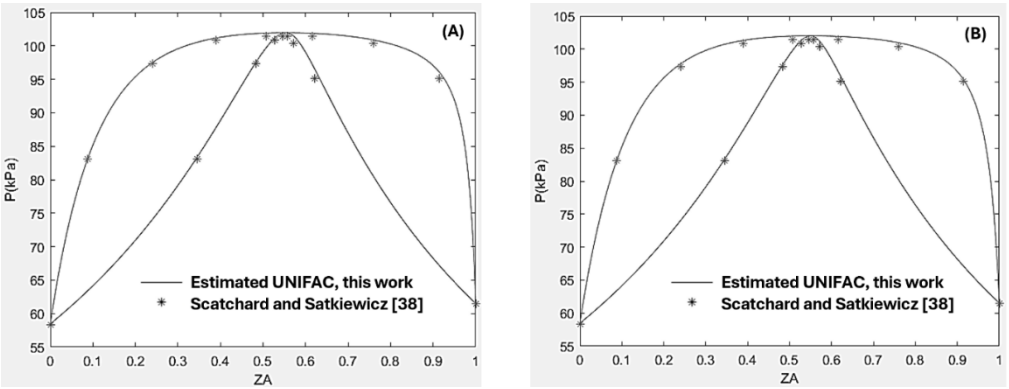

Figure 8 - Pressure–composition (*P*–*z*) diagram at 338 K comparing results obtained with group interaction parameters estimated using two different initial guesses: (A) the original DDBST values and (B) the manually modified DDBST values.

Despite the considerable numerical differences observed in the final parameter values resulting from the two initial guesses, the computed phase diagrams exhibit nearly indistinguishable behavior. This observation reinforces the idea that distinct parameter sets may converge to similarly accurate macroscopic predictions within the context of the UNIFAC model, even though their microscopic representations of group interactions differ substantially.

This behavior is also reflected in the AAD_P and AAD_Y values reported in Tables 15 and 16. In both cases, the deviations remain consistently low and close to each other, suggesting that both sets of optimized parameters are equally capable of capturing the thermodynamic behavior of the cyclohexane–ethanol system. Such convergence toward comparable modeling performance, despite the differences in parameter values and initial conditions, highlights the existence of multiple valid parameter sets that can adequately reproduce the experimental data.

7.5.2. COSMO-SAC-Based Parameter Optimization Strategy

7.5.2.1. Initial Guess: Original UNIFAC Interaction Parameters from the DDBST Database

In this section, the MATLAB optimization routine described in Section 6.4 (developed to incorporate activity coefficient data predicted by the COSMO-SAC model) is employed to estimate UNIFAC group interaction parameters. As with the previous methodology based on VLE data, two distinct initial guesses were used to evaluate the sensitivity of the optimization process to starting conditions.

The first initial guess consists of the original UNIFAC interaction parameters obtained directly from the DDBST (Dortmund Data Bank Software & Separation Technology) database, as listed in Table 1 (Section 6.2). The results obtained from successive optimization cycles using this starting point are presented in Table 17 below.

Table 17 - Estimated UNIFAC group interaction parameters after successive optimization cycles using the original DDBST parameter set as the initial guess. Optimizations were performed using MATLAB's *fminsearch* function (Nelder–Mead simplex algorithm), minimizing the deviation between UNIFAC-predicted and COSMO-SAC-predicted activity coefficients.

Parameters set	**Estimated UNIFAC Group Interaction Parameters**						**Objective Function**
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	986.5	-450.4	156.4	-817.7	-34.36	1913	28.84
1st Optimization	93967.28	-0.95	11823.82	197605.1	0.427	200100.6	28.53
2nd Optimization	93967.28	-0.95	11823.82	197605.1	0.427	200100.6	28.53

As observed in the table above, in contrast to the behavior previously seen in the VLE-based optimization routine (Section 7.5.1), the COSMO-SAC-based optimization exhibits rapid convergence to a specific parameter set. Notably, the objective function shows minimal reduction beyond the first iteration, and the estimated parameters remain stable in subsequent optimization cycles.

Moreover, unlike the results seen in earlier sections (where successive iterations often produced drastically different parameter values with marginal improvements in the objective function) here, the parameters do not undergo substantial fluctuations once the minimum is reached. This behavior may indicate the existence of a sharper or more well-defined minimum in the optimization landscape when using COSMO-SAC-derived activity coefficient data, as opposed to VLE experimental data.

Naturally, this outcome raises an important point regarding the sensitivity of the optimization results to the initial guess. It remains to be assessed whether similar convergence would occur if a different starting point were adopted, either toward the same parameter set or toward distinct, yet equally valid solutions, as observed in the VLE-based routine. This analysis is precisely the focus of the following section.

7.5.2.2. Optimization Initiated from a Parameter Set Previously Estimated Using Isothermal Experimental Data

In this second optimization scenario, the initial guess for the UNIFAC group interaction parameters was derived from a parameter set previously estimated in Section 7.3 of the present work. That parameter set had been specifically adjusted to reproduce isothermal experimental vapor-liquid equilibrium (VLE) data for the cyclohexane–ethanol binary mixture at six different temperatures: 278 K, 293 K, 308 K, 313 K, 323 K, and 338 K. The exact values employed as the initial guess in the current section are reported in Table 7.

The COSMO-SAC-based optimization routine, described in Section 6.4, was again applied here, this time using the isothermally derived parameters as the starting point. The results obtained over successive optimization cycles are presented in Table 18 below.

Table 18 - Estimated UNIFAC group interaction parameters after successive optimization cycles using as initial guess a parameter set previously estimated from isothermal VLE data (Section 3.3). Optimizations were carried out using MATLAB's *fminsearch* function (Nelder–Mead simplex algorithm), with the objective of minimizing the average absolute deviation between UNIFAC-predicted and COSMO-SAC-predicted activity coefficients.

Parameters set	**Estimated UNIFAC Group Interaction Parameters**						**Objective Function**
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	5881.4	-13.329	6923.4	344.24	75.069	1634.2	6.40
1st Optimization	6026.4	-10.45	836.67	34.91	157.32	1653.22	4.83
2nd Optimization	56209.14	14.11	1866.61	-120.32	238.19	1650.81	4.52
3rd Optimization	58164.53	14.11	1866.61	-120.32	238.19	1650.81	4.52

As observed in the previous case, the COSMO-SAC-based optimization procedure demonstrates a remarkably rapid convergence behavior compared to the VLE-based counterpart. After just a few iterations, the objective function stabilizes and the parameter values cease to exhibit the pronounced oscillations seen in the VLE-based routine. This is indicative of a more well-defined and narrow solution space when the activity coefficient profiles predicted by COSMO-SAC are used as the optimization target.

Another important observation emerges when comparing the outcome of this optimization to that obtained in the previous subsection (7.5.2.1), where the initial guess was based on the original UNIFAC parameters from the DDBST database. Although both runs employed the same COSMO-SAC-based optimization routine and objective function, they converged to markedly different sets of interaction parameters. This outcome suggests a higher sensitivity of the optimization landscape to the choice of initial guess than might be expected, as will be explored in detail in the next section.

7.5.2.3. Quantitative Analysis of COSMO-SAC-Based Optimizations Using Distinct Initial Guesses

In order to quantitatively assess the predictive capabilities of the final sets of UNIFAC group interaction parameters estimated using different starting points (namely, the original UNIFAC parameters from the DDBST database and the parameter set previously fitted to isothermal experimental data in Section 7.3) a comparative analysis was conducted using two performance metrics: the average absolute deviation in bubble pressure (AAD_P) and the average absolute deviation in the vapor-phase mole fraction of cyclohexane (AAD_Y).

These metrics were calculated at six distinct temperatures (278 K, 293 K, 308 K, 313 K, 323 K, and 338 K), using the final optimized parameter sets, i.e., those reported in the last row of Tables 17 and 18, respectively. The mathematical definitions of AAD_P and AAD_Y are provided in Equations (20) and (21) of the present work. The resulting values are summarized in Tables 19 and 20 below.

Table 19 - Average absolute deviations in bubble pressure (AAD_P) obtained using final group interaction parameters estimated with different initial guesses in the COSMO-SAC-based optimization routine.

Temperature (K)	AAD_P (kPa)	
	DDBST parameters as Initial Guess	*Parameters previously estimated with isothermal data as Inicial Guess*
278	0.213	0.128
293	0.331	0.180
308	0.365	0.321
313	0.669	0.271
323	0.578	0.690
338	3.124	1.621

Table 20 - Average absolute deviations in vapor-phase mole fraction of cyclohexane ($AAD_Y \times 10^{-2}$) obtained using final group interaction parameters estimated with different initial guesses in the COSMO-SAC-based optimization routine.

Temperature (K)	$AAD_Y \cdot 10^{-2}$	
	DDBST parameters as Initial Guess	*Parameters previously estimated with isothermal data as Inicial Guess*
278	1.820	0.659
293	1.489	0.507
308	0.485	0.734
313	0.960	0.601
323	0.820	0.488
338	1.018	0.529

As emphasized earlier, these results reflect the predictive accuracy of the final parameter sets obtained after full convergence of the COSMO-SAC-based optimization routine. Notably, although both initial guesses yielded parameter sets capable of modeling phase behavior to a certain extent, the magnitude of the deviations differs substantially depending on the starting point. In general, the parameter set derived from isothermal VLE data produced lower AAD_P and AAD_Y values across most temperatures, suggesting a superior alignment with the experimental vapor-liquid equilibrium behavior.

To further elucidate the practical implications of these differences, pressure–composition (P–z) diagrams were generated for both parameter sets across all six temperatures. Figures 9 and 10 display the results at the two limiting temperatures of 278 K and 338 K, while the intermediate cases (293 K, 308 K, 313 K, and 323 K) are provided in the Appendix (Figures A.5A–A.5H).

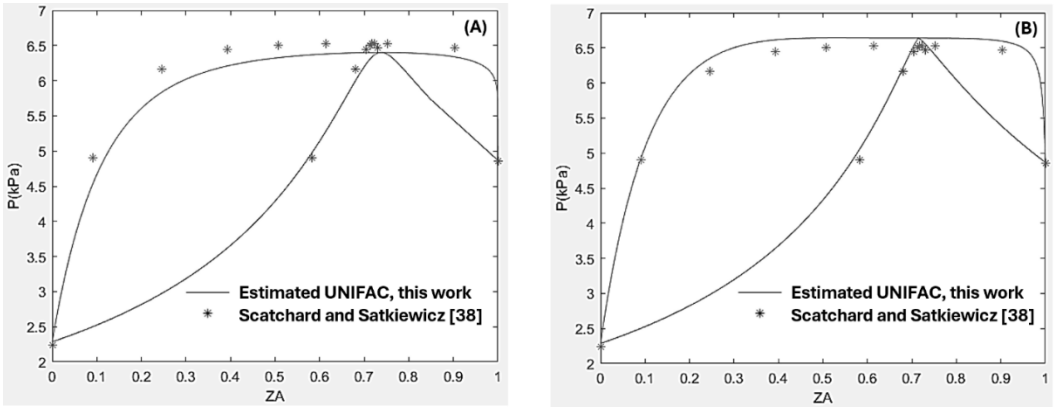

Figure 9 - Pressure–composition (*P–z*) diagram at 278 K comparing results obtained with group interaction parameters estimated using two different initial guesses: (A) the original DDBST values and (B) the parameter set previously fitted to isothermal VLE data.

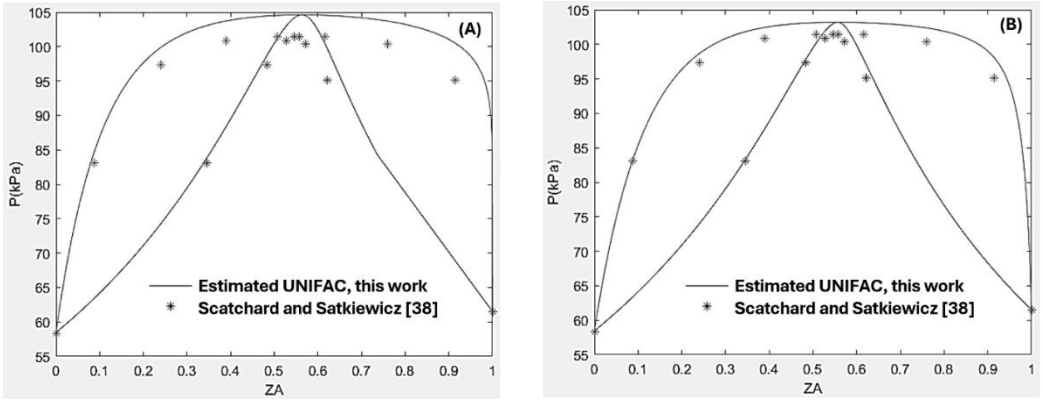

Figure 10 - Pressure–composition (*P–z*) diagram at 338 K comparing results obtained with group interaction parameters estimated using two different initial guesses: (A) the original DDBST values and (B) the parameter set previously fitted to isothermal VLE data.

Contrary to the findings from the ELV-based optimization routine (Section 7.5.1), in which distinct initial guesses led to different parameter sets that ultimately produced nearly indistinguishable phase diagrams, the COSMO-SAC-based routine demonstrates a markedly different behavior. Here, the use of different initial guesses leads not only to different optimized parameter sets, but also to noticeably distinct predictions of the phase behavior, clearly reflected in the divergence of the *P–z* curves in Figures 9 and 10.

This outcome further corroborates the findings presented in Tables 19 and 20, reinforcing the notion that the predictive performance of the final model is

significantly influenced by the choice of initial guess when COSMO-SAC-predicted activity coefficients are used as the optimization reference.

In summary, the COSMO-SAC-based optimization routine exhibits convergence to well-defined parameter sets that are stable across optimization cycles. However, unlike the ELV-based routine, the final parameter set and its predictive accuracy remain strongly dependent on the initial guess. While both routines are capable of exploring the parameter space effectively, the nature of the objective function in each case leads to fundamentally different optimization landscapes and convergence behaviors.

7.6. Evaluation of Optimization Outcomes Using COSMO-SAC Activity Coefficient Data at All Six Temperatures

Once the COSMO-SAC algorithm was fully integrated into the computational workflow, it became feasible to compute activity coefficients for both components of the cyclohexane–ethanol binary mixture across all six temperatures investigated throughout this work, namely 278 K, 293 K, 308 K, 313 K, 323 K, and 338 K. This capability enabled the execution of a more comprehensive optimization of the UNIFAC group interaction parameters, as described in Section 6.6 of the methodology, by incorporating thermodynamic data obtained from COSMO-SAC at all six temperature points.

A key objective of this stage was to explore whether the manner in which the activity coefficient vector is structured, i.e., the spatial distribution of the data points across the composition range, influences the outcomes of the optimization process or the predictive quality of the model. To this end, three distinct approaches were adopted in the construction of the activity coefficient vectors used as input for the optimization routines: (i) concentration of data points toward the region of higher cyclohexane mole fraction (i.e., the rich region), (ii) concentration of data points toward the region of lower cyclohexane mole fraction (i.e., the dilute region), and (iii) uniform spacing of data points across the entire composition range.

Each of these strategies was employed separately in the optimization procedure, allowing for a comparative analysis of their respective effects on the parameter estimation process and subsequent phase equilibrium predictions.

7.6.1. Optimization Using Activity Coefficient Data Concentrated in the Cyclohexane-Rich Region

In the first optimization scenario, the COSMO-SAC activity coefficient vectors were constructed so that the majority of the data points were concentrated near the upper end of the mole fraction scale for cyclohexane, i.e., in the region corresponding to more cyclohexane-rich mixtures.

As the initial guess for the optimization, the parameter set previously estimated from isothermal VLE data (presented in Section 7.3) was employed. This choice was justified based on the superior performance of that parameter set when used as an initial guess in comparative analyses conducted in Section 7.5.2.

The values of the initial parameter guess and the final optimized UNIFAC group interaction parameters are displayed in Table 21, along with the corresponding value of the objective function.

Table 21 - Estimated UNIFAC group interaction parameters using COSMO-SAC activity coefficient data concentrated in the cyclohexane-rich region. Optimization performed with initial guess from isothermal data (Section 7.3).

Parameters set	**Estimated UNIFAC Group Interaction Parameters**						**Objective Function**
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	5881.4	-13.329	6923.4	344.24	75.069	1634.2	14.02
Optimized Values	1202.39	13.16	8623.55	-115.51	234.64	1661.32	10.40

The MATLAB *fminsearch* routine was again used to minimize the objective function defined in Equation 19, which minimizes the relative deviation between COSMO-SAC activity coefficients and those calculated using the UNIFAC model. Although the initial guess used in this case was the same as that used in Table 18, the resulting value of the objective function is different because the optimization now considers data from six temperatures instead of only three.

In order to assess the predictive capability of the newly estimated parameters, vapor–liquid equilibrium (VLE) simulations were conducted using the optimized group interaction parameters. The average absolute deviations in bubble pressure (AAD_P) and in the vapor-phase mole fraction of cyclohexane (AAD_Y) were

calculated using Equations (23) and (24), respectively. These results are compiled in Table 22 below.

Table 22 - AAD_P and AAD_Y values for the COSMO-based routine using activity coefficient data concentrated in the cyclohexane-rich region.

Temperature (K)	AAD_P (kPa)	$AAD_y .10^{-2}$
278	0.124	0.629
293	0.172	0.485
308	0.307	0.711
313	0.296	0.585
323	0.665	0.521
338	1.558	0.586

The deviations observed in Table 22 are notably low across the entire temperature range, particularly at the lower temperatures. However, an increasing trend in AAD_P is observed at elevated temperatures, especially at 338 K, where the deviation exceeds 1.5 kPa. This behavior may be associated with increased vapor pressures and thermodynamic non-idealities at higher temperatures, which tend to exacerbate modeling inaccuracies.

Isothermal *P–z* phase diagrams were constructed to provide a visual assessment of the model's accuracy. The diagrams corresponding to the limiting temperatures of 278 K and 338 K are presented in Figure 11, while those for the intermediate temperatures (293 K, 308 K, 313 K, and 323 K) are provided in Appendix A (Figures A.6A–A.6D).

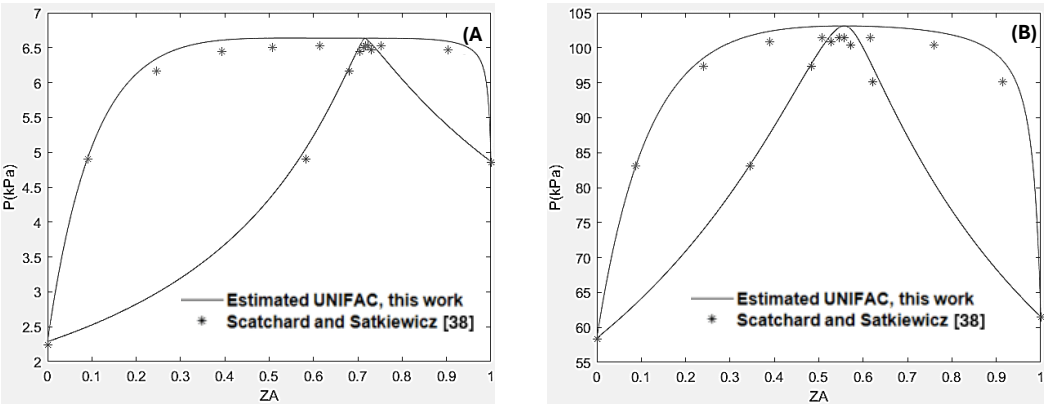

Figure 11 - Isothermal *P–z* Phase Diagrams for the Cyclohexane–Ethanol Binary System Using UNIFAC Parameters Estimated from COSMO-SAC Data Concentrated in the Cyclohexane-Rich Region: (A) 278 K and (B) 338 K.

7.6.2. Impact of Activity Coefficient Data Concentrated in the Dilute Cyclohexane Region on the Optimization Procedure Using the COSMO-SAC Model

In this second optimization experiment employing COSMO-SAC activity coefficient data across all six temperatures (278 K, 293 K, 308 K, 313 K, 323 K, and 338 K), a distinct strategy was adopted in the construction of the mole fraction grid used to generate the activity coefficient vectors. Specifically, the data points were intentionally concentrated toward the dilute region of cyclohexane, i.e., where the mole fraction of cyclohexane is low and the mixture is predominantly ethanol-rich.

As in the previous subsection (7.6.1), the initial guess for the group interaction parameters was the parameter set estimated from isothermal VLE data discussed in Section 7.3. The optimization itself was again carried out using MATLAB's *fminsearch* routine, which seeks to minimize the relative deviation between the COSMO-SAC activity coefficients and those predicted by the UNIFAC model, as defined by the objective function in Equation (19).

The resulting group interaction parameters obtained after optimization, along with the corresponding objective function value, are presented in Table 23. For comparison, the initial parameter values used as the starting point in the optimization process are also shown.

Table 23 - Estimated UNIFAC group interaction parameters using COSMO-SAC activity coefficient data concentrated in the cyclohexane-dilute region.

Parameters set	**Estimated UNIFAC Group Interaction Parameters**						**Objective Function**
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	5881.4	-13.329	6923.4	344.24	75.069	1634.2	4.61
Optimized Values	-3373.83	-66.88	18625.82	-66.66	302.27	-2141.45	0.81

The optimization procedure led to a marked reduction in the objective function value (from 4.61 to 0.81) indicating a significantly improved agreement between the UNIFAC model and the COSMO-SAC-generated activity coefficient data under this new distribution strategy. Notably, some of the optimized interaction parameters underwent large variations, including sign changes and substantial

increases in magnitude (e.g., a_{21} and a_{32}), which might reflect overfitting to the behavior in the dilute region.

To assess the capability of the optimized parameter set beyond the objective function minimization, isothermal vapor–liquid equilibrium calculations were performed, and average absolute deviations in bubble pressure (AAD_P) and vapor-phase mole fraction of cyclohexane (AAD_Y) were computed using Equations (23) and (24), respectively. These results are presented in Table 24.

Table 24 - AAD_P and AAD_Y values for the COSMO-based routine using activity coefficient data concentrated in the cyclohexane-dilute region.

Temperature (K)	AAD_P (kPa)	$AAD_y .10^{-2}$
278	0.133	1.094
293	0.186	0.889
308	0.304	0.915
313	0.418	0.894
323	0.508	0.596
338	1.098	0.445

A comparison between Table 24 (this subsection) and Table 22 (previous subsection) reveals that the strategy of concentrating activity coefficient data in the dilute region generally resulted in slightly inferior predictive performance compared to concentrating the data near the concentrated region of cyclohexane. This is particularly evident at temperatures between 278 K and 323 K, where AAD_P and AAD_Y values are consistently lower in the previous strategy. However, at the highest temperature (338 K), the current approach produced better predictive results, with a lower AAD_P (1.098 kPa versus 1.558 kPa) and AAD_Y (0.445% versus 0.586%).

This suggests that, for this binary system, thermodynamic behavior in the dilute region becomes more relevant at elevated temperatures, possibly due to increased volatility of ethanol, which dominates the composition in that range.

Isothermal vapor–liquid equilibrium phase diagrams were generated using the optimized group interaction parameters and are shown in Figure 12 for the two limiting temperatures (278 K and 338 K). The corresponding diagrams for the intermediate temperatures (293 K, 308 K, 313 K, and 323 K) are provided in Appendix A, Figures A.7A through A.7D.

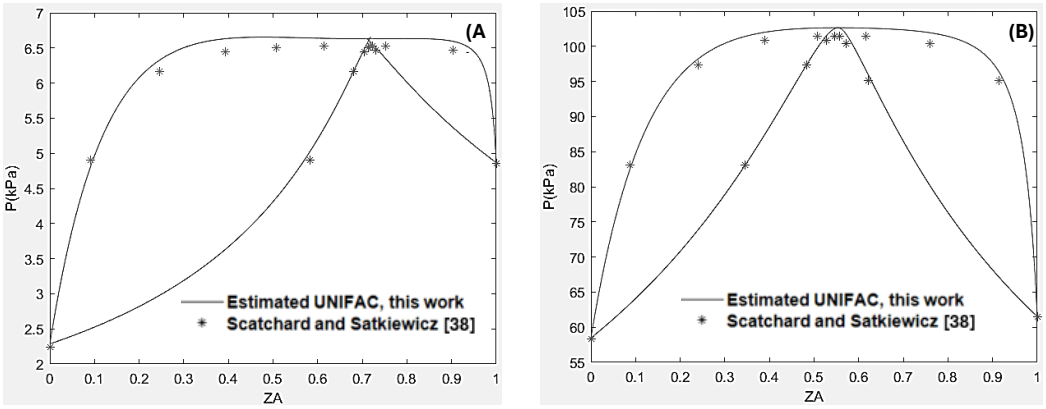

Figure 12 - Isothermal *P–z* Phase Diagrams for the Cyclohexane–Ethanol System Using UNIFAC Parameters Optimized with COSMO-SAC Data Concentrated in the Dilute Cyclohexane Region: (A) 278 K and (B) 338 K.

A visual inspection of the diagram at 278 K reveals an anomalous behavior in the bubble point curve: a local maximum appears prior to the azeotropic point, causing the dew point line to intersect and exceed the bubble point curve. This non-physical artifact, which is highlighted in Figure 13A with a zoomed-in view of the azeotropic region, suggests a deficiency in the model's performance under certain thermodynamic conditions.

Out of curiosity, this same region was examined in the previous optimization scenario (Section 7.6.1), where the activity coefficient data was concentrated in the cyclohexane-rich region. As shown in Figure 13B, the same anomalous behavior is present, although less pronounced, suggesting that this artifact may be a recurring feature of the model at lower temperatures, regardless of the data distribution used during parameter fitting.

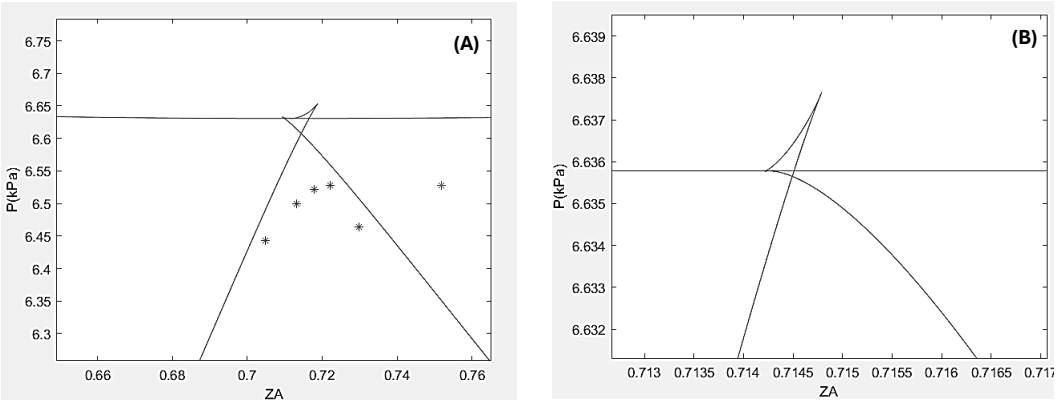

Figure 13 - Anomalous Azeotropic Behavior in the 278 K *P–z* Diagram: Dew Line Exceeds Bubble Line. (A) COSMO-SAC Data Concentrated in the Dilute Cyclohexane Region. (B) COSMO-SAC Data Concentrated in the Cyclohexane-Rich Region.

It is important to note that this irregularity appears only at low temperatures, specifically 278 K and, to a lesser extent, 293 K. At higher temperatures, such non-physical behaviors are not observed. In the next subsection (7.6.3), we examine whether using uniformly spaced activity coefficient data across the entire composition range produces more balanced results.

7.6.3. Optimization Results Using COSMO-SAC Activity Coefficient Data at All Six Temperatures — Case of Uniformly Spaced Composition Points

In this third optimization strategy based on COSMO-SAC activity coefficient data, values were sampled in a uniformly distributed manner across the entire composition range of the binary system composed of cyclohexane and ethanol. Specifically, activity coefficients were computed at evenly spaced mole fractions of cyclohexane, thus providing a balanced representation of the thermodynamic behavior of the mixture across both the dilute and concentrated regimes. As in the two previous cases, the initial guess for the group interaction parameters of the UNIFAC model corresponded to the parameter set previously estimated using isothermal vapor-liquid equilibrium (VLE) data, as presented in Section 7.3. This decision was based on the relatively strong performance observed when these parameters were employed in earlier optimization attempts (cf. Section 7.5.2).

The initial and optimized UNIFAC group interaction parameters, as well as the respective values of the objective function after the optimization process, are summarized in Table 25. The optimization was again performed using the

fminsearch function available in MATLAB, which employs the Nelder-Mead simplex algorithm.

Table 25 - Estimated UNIFAC Group Interaction Parameters Obtained from Optimization Using Uniformly Spaced COSMO-SAC Activity Coefficient Data Across Six Temperatures. Initial guess derived from isothermal VLE data (Section 7.3). Optimization performed with MATLAB's *fminsearch*.

Parameters set	**Estimated UNIFAC Group Interaction Parameters**						**Objective Function**
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	5881.4	-13.329	6923.4	344.24	75.069	1634.2	6.11
Optimized Values	-54.90	-62.34	59760000.00	-67.20	295.52	1092.80	1.47

It is important to highlight the exceptionally high value obtained for the a_{21} interaction parameter (on the order of 10^7), which may indicate numerical instability or overfitting during the optimization process. This behavior suggests that the model may be compensating for a lack of sensitivity in some regions of the composition space by inflating specific parameter values, thereby warranting further scrutiny or the incorporation of regularization techniques in future studies.

Following the optimization, average absolute deviations were computed to assess the accuracy of the model in reproducing VLE data. Specifically, the average absolute deviation in bubble pressure (AAD_P) and the average absolute deviation in the vapor-phase mole fraction of cyclohexane (AAD_Y) were calculated using Equations (23) and (24), respectively. The results are presented in Table 26.

Table 26 - Average Absolute Deviations in Bubble Pressure (AAD_P) and Cyclohexane Vapor-Phase Composition (AAD_Y) for the Uniformly Spaced Activity Coefficient Optimization Using COSMO-SAC Data at Six Temperatures.

Temperature (K)	**AAD_P (kPa)**	**$AAD_y .10^{-2}$**
278	0.130	1.047
293	0.186	0.849
308	0.314	0.889
313	0.372	0.866
323	0.579	0.601
338	1.309	0.497

The data from Table 26 suggest that this uniform spacing approach yields intermediate performance compared to the two previous strategies. While the AAD_Y values are moderately low and relatively consistent across temperatures, the AAD_P values, particularly at higher temperatures, are slightly elevated, indicating some loss of predictive accuracy for vapor pressure in those regions.

To further evaluate the model's performance, isothermal *P–z* phase diagrams were constructed for the lower and upper temperature boundaries, i.e., 278 K and 338 K, using the optimized group interaction parameters from Table 25. These diagrams are presented in Figure 14. The corresponding diagrams for intermediate temperatures (293 K, 308 K, 313 K, and 323 K) are provided in Appendix A (Figures A.8A–A.8D).

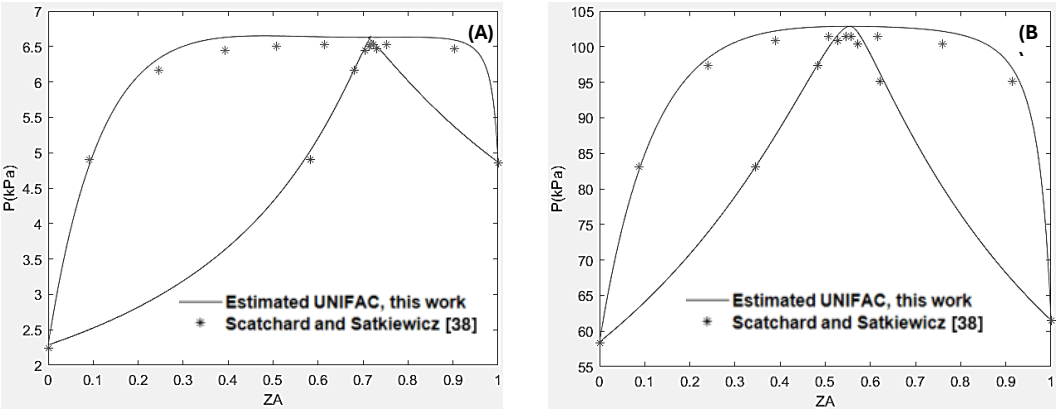

Figure 14 - Isothermal *P–z* Phase Diagrams for the Cyclohexane–Ethanol System at (A) 278 K and (B) 338 K. Calculations performed using UNIFAC parameters optimized with COSMO-SAC activity coefficients sampled uniformly across the composition range.

As can be seen in panel A of Figure 14, the vapor–liquid equilibrium curve at 278 K exhibits the same anomalous behavior previously observed in Figures 11 and 11. Specifically, the bubble point curve displays a local maximum prior to reaching the azeotropic composition. It is worth noting that this anomaly tends to appear only at lower temperatures (notably at 278 K and occasionally at 293 K), while disappearing entirely at elevated temperatures (308 K and above).

Finally, by comparing the results summarized in Tables 22, 24, and 26, as well as the diagrams in Figures 11, 12, and 14, it can be concluded that the choice of how the activity coefficient vectors are distributed (whether concentrated in a specific region of composition space or spread uniformly) does not dramatically affect the final quality of the optimization. All three strategies produce reasonably

accurate models, with marginal differences depending on temperature and region of composition. Nonetheless, additional insight may be gained by assessing the performance of these optimized parameters under isobaric conditions, which is the focus of the following section.

7.6.4. Comparative Evaluation of Activity Coefficient Vector Strategies in Isobaric VLE Predictions

In previous sections, it was observed that the choice of activity coefficient vector distribution (whether concentrated at high cyclohexane mole fractions, focused on the dilute region, or uniformly spaced) did not exert a significant influence on the quality of the optimization when applied to isothermal vapor–liquid equilibrium (VLE) data. Nevertheless, to further assess the robustness and generalizability of the parameter sets derived under these different configurations, it is essential to evaluate their predictive performance under isobaric conditions.

To this end, the UNIFAC group interaction parameters estimated at six isothermal temperatures (as reported in Tables 21, 23, and 25) were employed to compute isobaric VLE across five experimental pressure conditions for which binary phase equilibrium data are available: 40 kPa, 70 kPa, 98 kPa, 101.32 kPa, and 150 kPa.

The predictive accuracy of each parameter set was assessed using the average absolute deviation in bubble temperature (AAD_T), calculated according to Equation (22), and the average absolute deviation in vapor-phase mole fraction of cyclohexane (AAD_Y), as defined previously in Equation (21). The results are presented in Tables 27 and 28, respectively. Additionally, an illustrative comparison of the predicted isobaric T–z phase diagrams at atmospheric pressure (101.32 kPa) is provided in Figure 15, while the remaining diagrams for other pressures are included in Appendix A (Figures A.9A–A.11D).

Table 27 - Comparison of Average Absolute Deviations in Bubble Temperature (AAD_T) for Different Activity Coefficient Vector Configurations Used During Parameter Estimation. Isobaric VLE Predictions at Five Pressures.

Pressure (kPa)	**AAD_T (K)**		
	more concentrated region	*more dilute region*	*uniformly spaced*
40	1.21	0.88	0.94
70	1.75	1.38	1.47
98	2.03	1.65	1.76
101.32	1.67	1.4	1.47
150	2.28	2.09	2.18

Table 28 - Comparison of Average Absolute Deviations in Cyclohexane Vapor-Phase Composition (AAD_Y) for Different Activity Coefficient Vector Configurations Used During Parameter Estimation. Isobaric VLE Predictions at Five Pressures.

Pressure (kPa)	**$AAD_y .10^{-2}$**		
	more concentrated region	*more dilute region*	*uniformly spaced*
40	3.04	2.56	2.66
70	3.48	2.96	3.1
98	3.64	3.25	3.38
101.32	2.93	2.64	2.74
150	3.71	3.79	3.84

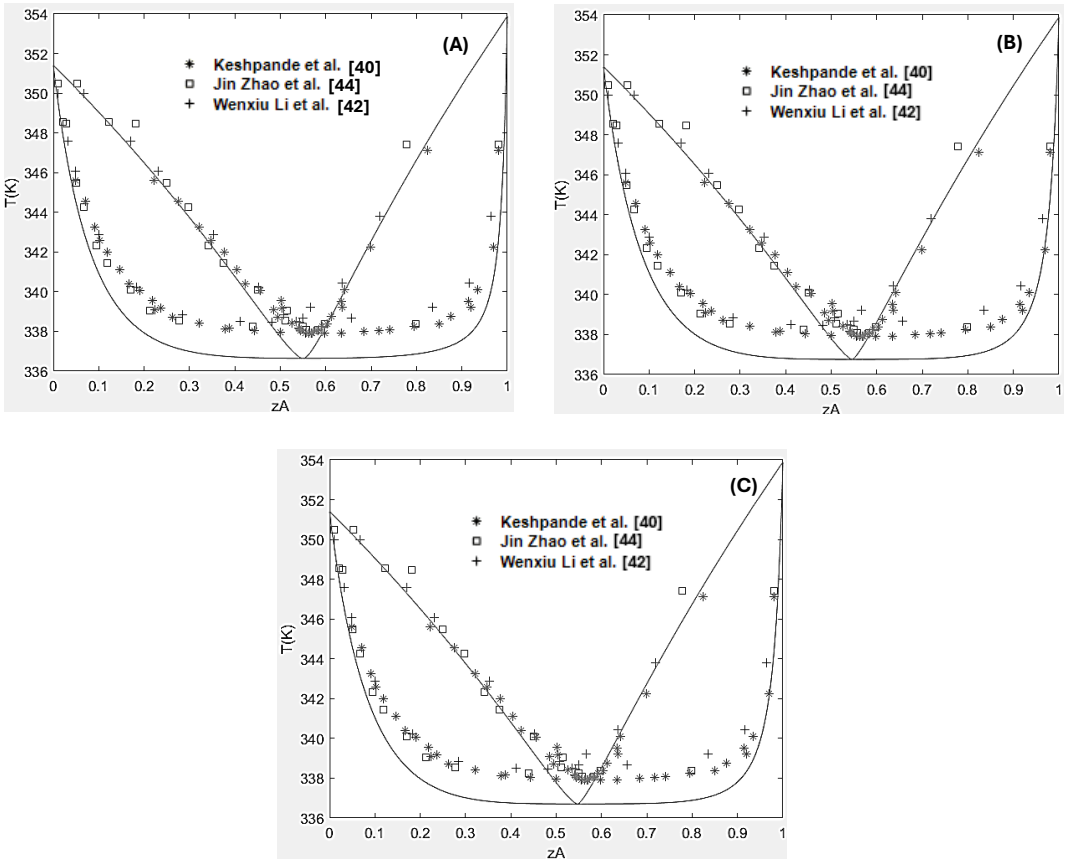

Figure 15 - Isobaric *T–z* Phase Diagrams for the Cyclohexane–Ethanol System at 101.32 kPa: Comparison of UNIFAC Parameter Sets Estimated Using Activity Coefficient Vectors (A) Concentrated Toward High Cyclohexane Mole Fractions, (B) Concentrated Toward the Dilute Region, and (C) Uniformly Spaced Across the Composition Range.

As evidenced by the numerical results in Tables 27 and 28 and the graphical profiles in Figure 15, the choice of distribution strategy for the activity coefficient vectors exerts only marginal influence on the quality of isobaric predictions. Differences among the three approaches are subtle, with the dilute-region-concentrated vector generally yielding slightly better performance in both AAD_T and AAD_Y across most pressure conditions. However, these improvements are not substantial enough to justify a preference based solely on isobaric prediction metrics.

This minimal sensitivity is further underscored by the isobaric phase diagrams shown in Figure 15. The visual differences among the three curves are nearly imperceptible, suggesting that the optimization landscape is relatively flat with respect to the distribution of activity coefficient vectors, at least within the range of temperatures and compositions considered here. This behavior is consistent with

the conclusions drawn from the isothermal optimization results presented in Section 7.6.3.

Nevertheless, it must be acknowledged that the absolute magnitudes of the deviations remain relatively high in the context of isobaric VLE modeling. When compared to the performance of the UNIFAC-DO model (see Tables 10 and 11 in Section 7.4), the optimized parameters derived from COSMO-SAC-informed activity coefficient vectors, regardless of their distribution strategy, fall short in reproducing isobaric phase behavior with equivalent accuracy.

This persistent discrepancy highlights the need for a more refined optimization approach. To address this challenge, a hybrid optimization strategy will be introduced in the following sections, with the objective of generating a parameter set capable of providing consistent and accurate predictions of phase equilibrium and activity coefficients.

7.7. Hybrid Parameter Estimation: Results and Discussion

7.7.1. First Optimization Cycle: Activity-Coefficient-Only Minimization

The first stage of the hybrid optimization strategy, as described in Section 6.7, consisted of estimating the UNIFAC group interaction parameters by minimizing the absolute deviations between activity coefficients of cyclohexane calculated with UNIFAC and those predicted by COSMO-SAC (Equation 20). The initial guess employed for this optimization cycle was derived from the parameter set previously estimated in Section 7.3 for isothermal VLE data, but this time each parameter was randomly perturbed by adding an integer between -300 and +300, generated with the MATLAB built-in function *randi*.

The estimated parameters are summarized in Table 29, which also includes the perturbed initial guess for comparison. The corresponding value of the objective function is reported in the last column. It is noteworthy that, unlike the optimizations described in Section 7.6.3, where some parameters diverged toward unrealistic magnitudes (e.g., a_{21} reaching the order of 10^7), all optimized values in this case remained within a physically reasonable range, on the order of 10^3 or less.

This outcome reflects the stabilizing effect of using absolute deviations as the error metric.

Table 29 - Estimated UNIFAC group interaction parameters obtained in the first cycle of the hybrid strategy. The objective function corresponds to the absolute deviation between activity coefficients of cyclohexane calculated with UNIFAC and those predicted by COSMO-SAC. The initial guess was generated from the parameters of Section 7.3 with random perturbations.

Parameters set	Estimated UNIFAC Group Interaction Parameters						Objective Function
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	6005.43	-294.33	6789.40	71.24	-166.93	1828.24	575.84
Optimized Values	1080.18	-78.28	5486.28	-52.91	309.29	1499.23	2.72

Following parameter estimation, the accuracy of the optimized model was further evaluated by computing the sum of absolute deviations of activity coefficients (SAD_γ) at each temperature according to Equation (26). In this formulation, the summation extends over 49 equidistant mole fraction points of cyclohexane, for which COSMO-SAC activity coefficients were calculated and subsequently compared with the corresponding UNIFAC predictions.

$$SAD_\gamma = \sum_{i=1}^{N} \left| \gamma_{i,CALC,j} - \gamma_{i,COSMO,j} \right| \tag{26}$$

where $\gamma_{i,CALC,j}$ and $\gamma_{i,COSMO,j}$ denote, respectively, for each component i, the activity coefficient predicted with the current UNIFAC parameter set and the activity coefficient obtained from the COSMO-SAC model for composition point j. The results are reported in Table 30, distinguishing the deviations for both cyclohexane and ethanol.

Table 30 - Sum of absolute deviations ($SAD\gamma$) between activity coefficients calculated with UNIFAC and those predicted by COSMO-SAC, for each component and at six temperatures.

Temperature (K)	SAD_γ	
	Cyclohexane	Ethanol
278	0.87	4.59
293	0.30	3.63
308	0.32	2.70
313	0.35	2.83
323	0.38	2.97
338	0.49	2.93

Two important observations can be made from these results. First, the SAD_γ values for cyclohexane are consistently below unity across all six temperatures, demonstrating that the optimization procedure was highly effective in minimizing deviations for this component, as expected since the objective function was constructed solely in terms of cyclohexane activity coefficients. Second, although ethanol was not directly included in the objective function, its activity coefficients were also reproduced with good accuracy, with SADγ values between 2.70 and 4.59 depending on the temperature. This outcome indicates that the Gibbs–Duhem constraint linking the activity coefficients of the two components was properly respected, such that improvements in cyclohexane predictions indirectly benefited ethanol as well.

7.7.2. Second-Cycle Optimization: Hybrid Refinement Incorporating VLE Data

The second stage of the hybrid optimization strategy employed the parameter set obtained in the first cycle (reported in Table 29) as the initial guess. This new optimization utilized the objective function defined in Equation (21), which simultaneously accounts for absolute deviations in activity coefficients, bubble pressures, and vapor-phase compositions of cyclohexane. By broadening the objective function to include experimental VLE data in addition to COSMO-SAC predictions, this stage aimed to refine the parameter set toward improved predictive performance across both thermodynamic properties.

The resulting optimized parameters are presented in Table 31, together with the initial guess values and the corresponding objective function. It should be emphasized that the absolute magnitude of the objective function at this stage is considerably larger than that obtained in the first cycle. This increase is expected, since the objective function now integrates deviations from multiple physical quantities rather than being restricted to activity coefficients alone. Notably, one of the parameters (a_{21}) exhibited a strong increase in magnitude, rising to the order of 10^6. While such behavior may reflect compensating adjustments within the nonlinear optimization landscape, it also highlights the sensitivity of the hybrid framework to particular interaction terms. Nevertheless, the overall objective function was substantially reduced compared to the initial guess, indicating a successful refinement of the parameter set.

Table 31 - Estimated UNIFAC group interaction parameters obtained in the second cycle of the hybrid optimization strategy. The objective function comprises absolute deviations in activity coefficients, bubble pressures, and vapor-phase mole fractions of cyclohexane.

Parameters set	**Estimated UNIFAC Group Interaction Parameters**						**Objective Function**
	a_{12}	a_{13}	a_{21}	a_{23}	a_{31}	a_{32}	
Initial Guess	1080.18	-78.28	5486.28	-52.91	309.29	1499.23	27794.72
Optimized Values	12337.10	45.79	5079550.13	-25.25	142.12	1544.44	7695.72

To evaluate the predictive performance of these parameters, isothermal and isobaric VLE calculations were carried out and compared with experimental data. Representative isothermal *P–z* diagrams are shown in Figure 16 for the lowest (278 K) and highest (338 K) experimental temperatures, while representative isobaric *T–z* diagrams are displayed in Figure 17 for the lowest (40 kPa) and highest (150 kPa) pressures. The remaining phase diagrams at intermediate conditions are provided in the Appendix. The visual inspection of Figures 16 and 17 reveals that the hybrid-optimized parameters successfully reproduce both the curvature and the azeotropic behavior of the cyclohexane–ethanol system over a wide range of conditions.

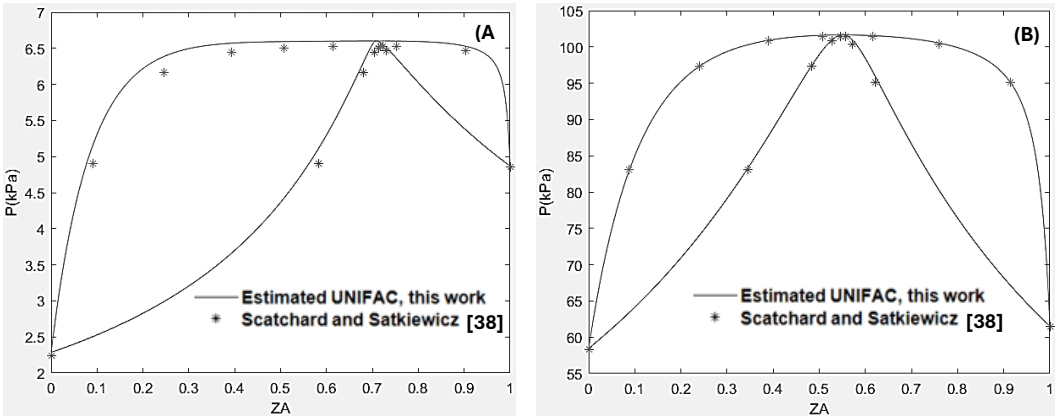

Figure 16 - Isothermal *P–z* phase diagrams for the cyclohexane–ethanol system at (A) 278 K and (B) 338 K. Predictions obtained using UNIFAC parameters optimized in the second cycle of the hybrid strategy.

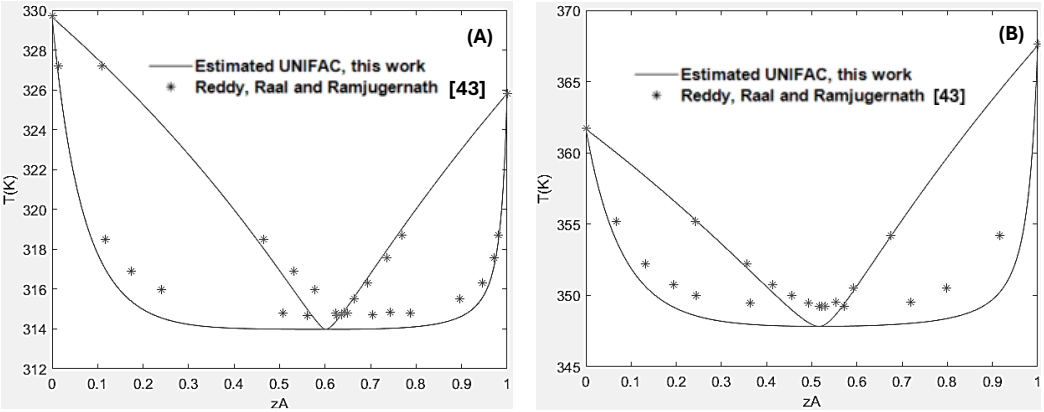

Figure 17 - Isobaric *T–z* phase diagrams for the cyclohexane–ethanol system at (A) 40 kPa and (B) 150 kPa. Predictions obtained using UNIFAC parameters optimized in the second cycle of the hybrid strategy.

The quantitative predictive accuracy was assessed in terms of the average absolute deviations. Specifically, the average absolute deviation in bubble pressure (AAD_P) and in vapor-phase composition of cyclohexane (AAD_Y) were calculated for isothermal simulations according to Equations (23) and (24), while the average absolute deviation in bubble temperature (AAD_T) and AAD_Y were evaluated for isobaric simulations using Equation (25). The results are summarized in Tables 32 and 30.

Table 32 - Average absolute deviations in bubble pressure (AAD_P) and vapor-phase mole fraction of cyclohexane (AAD_Y) for isothermal VLE predictions using parameters obtained in the second cycle of the hybrid strategy.

Temperature (K)	AAD_P (kPa)	$AAD_y .10^{-2}$
278	0.13	0.72
293	0.13	0.47
308	0.20	1.00
313	0.49	0.31
323	0.04	0.20
338	0.09	0.30

Table 33 - Average absolute deviations in bubble temperature (AAD_T) and vapor-phase mole fraction of cyclohexane (AAD_Y) for isobaric VLE predictions using parameters obtained in the second cycle of the hybrid strategy.

Pressure (kPa)	AAD_T (K)	$AAD_y .10^{-2}$
40	0.84	2.57
70	1.26	2.86
98	1.45	2.91
101.32	1.24	2.51
150	1.78	3.23

The numerical results presented in Tables 32 and 33 highlight the effectiveness of the hybrid approach in reconciling COSMO-SAC predictions with experimental VLE data. For isothermal simulations, AAD_P values remain consistently below 0.5 kPa, and AAD_Y values do not exceed 1.0×10^{-2} for most temperatures, demonstrating strong consistency between predicted and experimental bubble pressures as well as vapor-phase compositions. Similarly, for isobaric simulations, AAD_T values remain within 2 K even at the highest pressure (150 kPa), while AAD_Y values remain below 3.3×10^{-2}.

Overall, the second cycle of optimization yielded a parameter set capable of providing accurate predictions for both isothermal and isobaric VLE conditions. Nevertheless, to assess its robustness and practical applicability, it is necessary to compare these results against those obtained with the UNIFAC-Do model and with the parameters estimated exclusively from isothermal VLE data (Section 7.3). This comparative evaluation is the focus of the next subsection.

7.7.3. Comparative Evaluation of UNIFAC-DO, Isothermal Fitted, and Hybrid Optimized Parameters

Having completed the optimization procedures described in the previous sections, it is now possible to perform a comprehensive comparison among three distinct parameter sets: (i) the original UNIFAC-DO parameters, (ii) the parameters estimated exclusively from isothermal VLE data (Section 7.3), and (iii) the parameters obtained via the hybrid optimization strategy introduced in Section 7.7.2. It is important to recall that the UNIFAC-DO model employs 18 group interaction parameters, whereas both parameter sets developed in this work rely on only six.

The comparative analysis was carried out in three stages. First, the ability of each parameter set to reproduce activity coefficients for both cyclohexane and ethanol was assessed. Second, the parameter sets were compared with respect to their predictive accuracy in isothermal VLE simulations. Finally, their performance was evaluated in isobaric VLE calculations.

Tables 34 and 35 present the sum of absolute deviations (SAD_γ) in activity coefficients for cyclohexane and ethanol, respectively, across the six investigated temperatures.

Table 34 - Sum of absolute deviations in activity coefficients ($SAD\gamma$) for cyclohexane: comparison between UNIFAC-DO, isothermal-fitted, and hybrid-optimized parameters.

Temperature (K)	SAD_γ for Cyclohexane		
	UNIFAC-DO	Isothermal Fitted Parameters	Hybrid Approach
278	6.63	7.956	8.75
293	5.29	6.359	6.58
308	4.71	5.045	4.96
313	4.69	4.659	4.53
323	5.06	3.971	3.87
338	6.54	3.143	3.33

Table 35 - Sum of absolute deviations in activity coefficients ($SAD\gamma$) for ethanol: comparison between UNIFAC-DO, isothermal-fitted, and hybrid-optimized parameters.

Temperature (K)	SAD_γ for Ethanol		
	UNIFAC-DO	Isothermal Fitted Parameters	Hybrid Approach
278	11.63	6.842	3.92
293	8.07	4.533	3.93
308	7.71	3.240	4.47
313	8.18	3.018	4.74
323	9.80	2.777	5.35
338	13.34	2.625	6.39

From Table 34, it can be observed that UNIFAC-DO generally provides lower deviations for cyclohexane at lower temperatures (278–308 K). However, the hybrid parameters outperform both alternatives at 313 K and 323 K, while the isothermal-fitted parameters yield the best result at 338 K. In contrast, Table 35 reveals a different pattern for ethanol: UNIFAC-DO consistently produces the

largest deviations across all temperatures, highlighting its limitations in describing ethanol activity coefficients. Here, the hybrid parameters show superior performance at the lowest temperatures (278 K and 293 K), whereas the isothermal-fitted set is more effective at intermediate and higher temperatures. These results emphasize that while cyclohexane predictions remain challenging, significant improvements over UNIFAC-DO can be achieved for ethanol by adopting reduced-parameter models.

The second stage of comparison involved isothermal VLE calculations, where predictive accuracy was evaluated in terms of average absolute deviations in bubble pressure (AAD_P) and vapor-phase mole fraction of cyclohexane (AAD_Y). Results are summarized in Tables 36 and 37.

Table 36 - Average absolute deviations in bubble pressure (AAD_P) for isothermal VLE predictions.

Temperature (K)	AAD_P (kPa)		
	UNIFAC-DO	Isothermal Fitted Parameters	Hybrid Approach
278	0.11	0.04	0.13
293	0.25	0.06	0.13
308	0.48	0.18	0.20
313	1.21	0.52	0.49
323	1.07	0.11	0.04
338	2.11	0.83	0.09

Table 37 - Average absolute deviations in vapor-phase mole fraction of cyclohexane (AAD_Y) for isothermal VLE predictions.

Temperature (K)	$AAD_y .10^{-2}$		
	UNIFAC-DO	Isothermal Fitted Parameters	Hybrid Approach
278	1.01	0.56	0.72
293	0.86	0.30	0.47
308	0.89	0.83	1.00
313	0.93	0.21	0.31
323	0.84	0.18	0.20
338	1.01	0.20	0.30

These results clearly demonstrate the strength of the isothermal-fitted parameters under the conditions they were specifically optimized for: in both AAD_P and AAD_Y, this set consistently outperforms the other two at most temperatures.

However, the hybrid approach shows a marked advantage at higher temperatures (323 K and 338 K), particularly in reproducing bubble pressures, where its deviations are as low as 0.04 kPa and 0.09 kPa, respectively.

Finally, Tables 38 and 39 present the performance of the three parameter sets in reproducing isobaric VLE data, measured by the average absolute deviations in bubble temperature (AAD_T) and vapor-phase mole fraction of cyclohexane (AAD_Y).

Table 38 - Average absolute deviations in bubble temperature (AAD_T) for isobaric VLE predictions using UNIFAC-DO, isothermal-fitted, and hybrid-optimized parameters.

Pressure (kPa)	AAD_T (K)		
	UNIFAC-DO	Isothermal Fitted Parameters	Hybrid Approach
40	0.43	0.93	0.84
70	0.69	1.48	1.26
98	0.72	1.73	1.45
101.32	0.65	1.48	1.24
150	0.98	2.07	1.78

Table 39 - Average absolute deviations in the vapor-phase mole fraction of cyclohexane (AAD_Y) for isobaric VLE predictions using UNIFAC-DO, isothermal-fitted, and hybrid-optimized parameters.

Pressure (kPa)	$AAD_Y .10^{-2}$		
	UNIFAC-DO	Isothermal Fitted Parameters	Hybrid Approach
40	1.79	2.7	2.57
70	1.81	3.17	2.86
98	1.82	3.27	2.91
101.32	1.8	2.81	2.51
150	2.36	3.49	3.23

As shown in Tables 38 and 39, the UNIFAC-DO parameter set consistently outperforms both the isothermal-fitted and hybrid-optimized sets in reproducing isobaric VLE data. Across all five pressure conditions, UNIFAC-DO yields lower deviations in both bubble temperature and vapor-phase composition, demonstrating a clear advantage under isobaric conditions. For example, at 98 kPa, the AAD_T obtained with UNIFAC-DO is only 0.72 K, whereas the deviations increase to 1.73 K and 1.45 K for the isothermal-fitted and hybrid sets, respectively. Similarly, for

AAD_Y at 150 kPa, the UNIFAC-DO model achieves 2.36×10^{-2}, compared to 3.49×10^{-2} for the isothermal-fitted parameters and 3.23×10^{-2} for the hybrid parameters.

One possible explanation for this behavior is that the original UNIFAC-DO parameterization may have been explicitly or implicitly tuned to reproduce isobaric equilibrium data during its development. If this is the case, the superior performance of UNIFAC-DO under isobaric conditions would reflect a bias in the original parameter optimization strategy, favoring constant-pressure data at the expense of broader transferability.

Nevertheless, although the optimized parameter sets developed in this work do not surpass UNIFAC-DO in isobaric VLE predictions, their performance remains within acceptable limits. Deviations below 2 K in bubble temperature and below 3.5×10^{-2} in vapor-phase composition indicate that both the isothermal-fitted and hybrid approaches can provide competitive predictive accuracy, particularly given that they rely on only six group interaction parameters rather than the 18 parameters employed in UNIFAC-DO.

When all comparisons are jointly considered (namely, the reproduction of activity coefficients, isothermal VLE, and isobaric VLE) the hybrid optimization approach emerges as perhaps the most balanced strategy among the three alternatives. In isobaric conditions, it consistently outperforms the isothermal-fitted parameters across all investigated pressures (Tables 38 and 39), although its accuracy does not reach the level achieved by the original UNIFAC-DO model. Under isothermal conditions, on the other hand, the hybrid parameters surpass UNIFAC-DO at nearly all temperatures examined (Tables 36 and 37), while generally performing slightly worse than the isothermal-fitted parameters (except at 313 K, 323 K, and 338 K), where the hybrid set achieves superior agreement in bubble pressure predictions (AAD_P), even exceeding the performance of the isothermal-fitted parameters. Finally, in terms of activity coefficient reproduction, the UNIFAC-DO model performs significantly worse, while the isothermal-fitted and hybrid parameters display a complementary behavior, with one or the other showing superior accuracy depending on the temperature.

8. Conclusions and Future Works

The present work investigated different strategies for the estimation and optimization of thermodynamic model parameters, with particular emphasis on the prediction of pure component properties and the description of non-ideal phase equilibria in the cyclohexane + ethanol system. Initially, the estimation of the adjustable parameters (λ_1, λ_2, λ_3) of the Almeida–Aznar–Telles cohesion function in the Peng–Robinson equation of state provided satisfactory reproduction of the vapor pressure curves for both pure compounds. The obtained correlations captured the temperature dependence of the saturation pressure with average absolute deviations of 0.091 bar for cyclohexane and 0.076 for ethanol, indicating a good agreement with the saturation pressures obtained from the REFPROP software.

The UNIFAC-Do model was subsequently applied to describe the binary system, with its performance assessed under both isothermal and isobaric conditions. The results highlighted the capability of this model to represent the VLE of the system, particularly under isobaric conditions. Deviations in bubble pressure predictions were below 2.11 kPa, while deviations in bubble temperature predictions were below 0.98 K. These findings confirm the robustness of the UNIFAC-Do approach as a predictive tool, although 18 parameters are required for its application.

When parameters were adjusted specifically to isothermal VLE data, the predictive accuracy improved under the same conditions, as expected. In isothermal representations, average deviations were below 0.83 kPa, showing better alignment with experimental reference data. However, when these same parameters were applied to isobaric conditions, their transferability was limited, as deviations in saturation pressure increased beyond those obtained with UNIFAC-Do. This outcome reinforces the trade-off between local optimization for a given type of data and generalization across different operating conditions.

The influence of the initial guess in the optimization routines was also systematically examined. Optimizations conducted with the *fminsearch* algorithm

revealed that different starting points converge to distinct local minima, resulting in variations in the estimated interaction parameters. While certain initializations, such as those derived from modified DDBST parameters, yielded improved convergence behavior and lower objective function values, others led to suboptimal solutions. This analysis underscores the importance of carefully selecting initial estimates in nonlinear parameter regressions to avoid biased or physically unrealistic results.

The use of alternative distributions of activity coefficient data, derived from COSMO-SAC and weighted towards cyclohexane-rich, ethanol-rich, or uniformly spaced regions, was tested in order to verify their influence on the optimization outcome. The results indicated that the choice of data distribution exerted only marginal influence on the quality of both isothermal and isobaric predictions, suggesting that the optimization procedure is not particularly sensitive to this factor, at least within the studied system and methodological framework.

Finally, a hybrid optimization strategy combining two successive optimization cycles was explored. This approach emerged as the most balanced among the alternatives considered, since it surpassed the isothermal-fitted parameters under isobaric conditions and outperformed the UNIFAC-Do model in most isothermal cases. Although not always the most accurate in every individual comparison, the hybrid set consistently delivered a reliable compromise, providing complementary behavior in reproducing activity coefficients and phase-equilibrium data. These results highlight the value of hybrid strategies as a practical middle ground between activity-coefficient prediction and VLE prediction, both isothermal and isobaric, and suggest their broader applicability to other complex non-ideal systems.

As a natural continuation of the present work, several directions for future research can be identified. First, the construction of a more robust optimization framework, combining global and local search techniques, is recommended to mitigate the sensitivity of the regression results to initial parameter guesses. Additionally, the implementation of the φ–φ approach, either through the adoption of Gibbs excess mixing rules coupled to the Peng–Robinson equation of state or by employing group-contribution lattice-fluid equations of state, may provide valuable alternatives for describing highly non-ideal systems. Finally, extending the methodology to predict VLE data for other binary systems involving similar

functional groups, followed by systematic comparisons against UNIFAC-Do predictions available in the Dortmund Data Bank, would serve both to validate the proposed strategies and to assess their transferability to a broader range of mixtures.

9. References

1. BARRIENTOS, D. A. et al. Recent Advances in Reactive Distillation. **Engineering Proceedings**, 56, n. 99, 2023.

2. PRAUSNITZ, J. M.; LICHTENTHALER, R. N.; DE AZEVEDO, E. G. **Molecular Thermodynamics of Fluid-Phase Equilibria**. 3rd. ed. Upper Saddle River, NJ: Prentice Hall, 1999.

3. REID, R. C.; PRAUSNITZ, J. M.; POLING, B. E. **The Properties of Gases and Liquids**. 4th. ed. New York: McGraw-Hill, 1987.

4. ELLIOTT, J. R.; LIRA, C. T. **Introductory Chemical Engineering Thermodynamics**. 2nd. ed. Upper Saddle River, NJ: Prentice Hall, 1999.

5. HSIEH, C.-M.; SANDLER, S. I.; LIN, S.-T. Improvements of COSMO-SAC for vapor–liquid and liquid–liquid equilibrium predictions. **Fluid Phase Equilibria**, 2010. 90-97.

6. SMITH, J. M.; VAN NESS, H. C.; ABBOTT, M. M. **Introduction to Chemical Engineering Thermodynamics**. 7th ed. ed. [S.l.]: McGraw-Hill, 2005.

7. GMEHLING, J.; RASMUSSEN, P. Flash Points of Flammable Liquid Mixtures Using UNIFAC. **Industrial & Engineering Chemistry Fundamentals**, 21, 1982. 186-188.

8. WANG, X.; CHENG, B. Integrating molecular dynamics simulations and experimental data for azeotrope predictions in binary mixtures. **The Journal of Chemical Physics**, 15 July 2024.

9. KISTER, H. Z. **Distillation Design**. 1st. ed. New York: McGraw-Hill, 1992.

10. TAYLOR, R.; KRISHNA, R. **Multicomponent Mass Transfer**. 1t. ed. New York: Wiley, 1993.

11. YE LI et al. Process Assessment of Heterogeneous Azeotropic Dividing-Wall Column for Ethanol Dehydration with Cyclohexane as an Entrainer: Design and Control. **Industrial & Engineering Chemistry Research**, 55, n. 32, 2016. 8784-8801.

12. MIRANDA N.T.; MACIEL FILHO R.; WOLF MACIEL M.R. Comparison of Complete Extractive and Azeotropic Distillation Processes for Anhydrous Ethanol Production Using Aspen Plus™ Simulator. **Chemical Engineering Transactions**, 80, 2020. 43-48.

13. FREDENSLUND, A.; JONES, R. L.; PRAUSNITZ, J. M. Group-Contribution Estimation of Activity Coefficients in Nonideal Liquid Mixtures. **AIChE Journal**, November 1975. 1086.

14. PRIVAT, R.; JAUBERT, J.-N. Thermodynamic Models for the Prediction of Petroleum-Fluid Phase Behaviour. In: _____ **Crude Oil Emulsions - Composition Stability and Characterization**. Rijeka: InTech, 2012. p. 71-107.

15. ZAVALA, M. S.; AROCHE, F. B.; BAZÚA, E. R. Comparative study of mixing rules for cubic equations of state in the prediction of multicomponent vapor-liquid equilibria. **Fluid Phase Equilibria**, 122, 1996. 99-116.

16. CHABAB, S.; COQUELET, C.; RIVOLLET, F. Generalization of the Wong-Sandler mixing rule to a generic cubic equation of state: Examples of use for systems of industrial interest (Hydrogen, CCUS, refrigeration). **The Journal of Supercritical Fluids**, 212, 2024.

17. LOPEZ-ECHEVERRY, J. S.; REIF-ACHERMAN, S.; ARAUJO-LOPEZ, E. Peng-Robinson equation of state: 40 years through cubics. **Fluid Phase Equilibria**, 2017. 39-71.

18. PENG, D.-Y.; ROBINSON, D. B. A New Two-Constant Equation of State. **Industrial & Engineering Chemistry Fundamentals**, 15, n. 1, 1976. 59-64.

19. YOUNG, A. F.; PESSOA, F. L. P.; AHÓN, V. R. R. Comparison of 20 Alpha Functions Applied in the Peng–Robinson Equation of State for Vapor Pressure Estimation. **Industrial & Engineering Chemistry Research**, 55, n. 22, 2016. 6506-6516.

20. YOUNG, A. F.; PESSOA, F. L. P.; AHÓN, V. R. R. Comparison of volume translation and co-volume functions applied in the Peng-Robinson EoS for volumetric corrections. **Fluid Phase Equilibria**, 435, 2017. 73-87.

21. AZNAR, M.; TELLES, A. S. A data bank of parameters for the attractive coefficient of the Peng-Robinson equation of state. **Brazilian Journal of Chemical Engineering**, 14, n. 1, 1997.

22. ALMEIDA, G. S.; AZNAR, M.; TELLES, A. S. Uma Nova Forma de Dependência com a Temperatura do Termo Atrativo de Equações de Estado Cúbicas. **Cadernos de Engenharia Química**, 8, 1991. 95-123.

23. PRIVAT, R.; JAUBERT, J.-N. The state of the art of cubic equations of state with temperature-dependent binary interaction coefficients: From correlation to prediction. **Fluid Phase Equilibria**, 567, 2023. 113697.

24. HURON, M.-J.; VIDAL, J. New mixing rules in simple equations of state for representing vapour-liquid equilibria of strongly non-ideal mixtures. **Fluid Phase Equilibria**, 3, n. 4, 1979. 255-271.

25. JAUBERT, J.-N.; MUTELET, F. VLE predictions with the Peng–Robinson equation of state and temperature dependent kij calculated

through a group contribution method. **Fluid Phase Equilibria**, 224, n. 2, 2004. 285-304.

26. WIENKE, G.; GMEHLING, J. Prediction of Octanol-Water Partition Coefficients, Henry Coefficients and Water Solubilities Using UNIFAC. **Toxicological & Environmental Chemistry**, 65, 1998. 57-86.

27. FREDENSLUND, A.; GMEHLING, J.; RASMUSSEN, P. Vapor-Liquid Equilibria Using UNIFAC, A Group Contribution Method. Amsterdam: Elsevier, 1997.

28. MICHELSEN, M. L.; MOLLERUP, J. **Thermodynamic Models:** Fundamentals & Computational Aspects. 2nd. ed. Holte: Tie-line Publications, 2007.

29. SANDLER, S. I. **Chemical and Engineering Thermodynamics**. 4th. ed. Hoboken, NJ: John Wiley & Sons, Inc., 2006.

30. DORTMUND Data Bank. Disponivel em: <https://www.ddbst.com/>. Acesso em: 29 November 2024.

31. PARAMETERS of the Modified UNIFAC (Dortmund) Model. **DDBST**. Disponivel em: <https://www.ddbst.com/PublishedParametersUNIFACDO.html>. Acesso em: 21 September 2024.

32. BELL, I. H. et al. A Benchmark Open-Source Implementation of COSMO-SAC. **Journal of Chemical Theory and Computation**, 16, 14 February 2020. 2635−2646. Disponivel em: <https://doi.org/10.1021/acs.jctc.9b01016>.

33. KLAMT, A. et al. Refinement and Parametrization of COSMO-RS. **The Journal of Physical Chemistry**, 5 June 1998. 5074-5085.

34. FERRARINI, F. et al. An open and extensible sigma-profile database for COSMO-based models. **AIChE Journal**, 27 April 2018. 3443-3455.

35. MULLINS, E. et al. Sigma-Profile Database for Using COSMO-Based Thermodynamic Methods. **Industrial & Engineering Chemistry Research**, 45, May 2006.

36. SCHULZE-HULBE, A. et al. n-Octane with Branched Pentanol Isomers: Isobaric Binary Vapor−Liquid Equilibria Measurements and Thermodynamic Modeling with NRTL, UNIFAC, and Modified UNIFAC (Dortmund). **Journal of Chemical & Engineering Data**, 70, 13 January 2025. 1083−1095.

37. TSUJI, T. et al. Vapor-liquid equilibrium (VLE) measurements of ethanol –heptane at isothermal (363.15, 393.15 and 423.15 K) and isobaric (101.33 kPa) conditions and correlation of liquid viscosity data. **Chemical Thermodynamics and Thermal Analysis**, 6, June 2022.

38. SCATCHARD, G.; SATKIEWICZ, F. G. Vapor-Liquid Equilibrium. XII. The System Ethanol-Cyclohexane from 5 to 65°. **Journal of the American Chemical Society**, 1964. 130-133.

39. SHUJIE LI et al. Vapor-liquid equilibrium in the binary and ternary systems containing ethyl propionate, ethanol and alkane at 101.3 kPa. **Fluid Phase Equilibria**, 476, 2018.

40. KUAN SHIH YUAN et al. Vapor-Liquid Equilibria. **Journal of Chemical & Engineering Data**, 9, 1963. 157.

41. JOSEPH, M. A.; RAMJUGERNATH, D.; RAAL, J. D. Computer-aided Measurement of Vapour-Liquid Equilibria in a Dynamic Still at Sub-Atmospheric Pressures. **Developments in Chemical Engineering and Mineral Processing**, 2002.

42. WENXIU LI et al. Measurement and correlation of the vapor-liquid equilibrium of cyclohexane + ethanol containing ILs and the thermophysical properties of these ILs at 101.3 kPa. **The Journal of Chemical Thermodynamics**, 150, 2020. 106203.

43. REDDY, P.; RAAL, J. D.; RAMJUGERNATH, D. A novel dynamic recirculating apparatus for vapour–liquid equilibrium measurements at moderate pressures and temperatures. **Fluid Phase Equilibria**, n. 358, 2013.

44. JIN ZHAO et al. Isobaric vapor–liquid equilibria for ethanol–water system containing different ionic liquids at atmospheric pressure. **Fluid Phase Equilibria**, 242, 2006.

45. XU XIN-HUA; WANG XIAO-GANG; LIU MEI-CHUAN. New experimental method for determination of binary liquid——gas phase diagram. **LABORATORY SCIENCE**, 18, 2015.

46. HAAZ, E. et al. Vapor−Liquid Equilibrium Study of the Monochlorobenzene−4,6-Dichloropyrimidine Binary System. **ACS Omega**, 7, 19 May 2022. 17670–17678.

47. YOUNG, A. F. et al. **Modelagem do equilíbrio líquido-vapor de sistemas binários por metodologia de Heidemann-Kokal:** aplicação em sistemas simples. IX Congresso Brasileiro de Termodinâmica Aplicada e V Escola de Termodinâmica. Porto Alegre-RS: [s.n.]. 2017.

48. SANDLER, S. I. **Chemical, biochemical, and engineering thermodynamics**. 5ª. ed. Hoboken: Wiley, 2017.

49. NELDER, J. A.; MEAD, R. A simplex method for function minimization. **The Computer Journal**, 7, January 1965. 308–313.

50. LAGARIAS, J. C. et al. Convergence properties of the Nelder-Mead simplex method in low dimensions. **Society for Industrial and Applied Mathematics**, 9, 2 December 1998. 112 147.

51. NATIONAL Institute of Standards and Technology. NIST Chemistry WebBook. Disponivel em: <https://webbook.nist.gov/>. Acesso em: 29 November 2024.

52. LEMMON, E. W. et al. NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 10.0, National Institute of Standards and Technology, Gaithersburg, 2018. Disponivel em: <https://www.nist.gov/srd/refprop>.

53. XIONG, R.; SANDLER, S. I.; BURNETT, R. I. An Improvement to COSMO-SAC for Predicting Thermodynamic Properties. **Industrial & Engineering Chemistry Research**, 53, 8 April 2014. 8265-8278. Disponivel em: <https://doi.org/10.1021/ie404410v>.

54. ZHOU, Y. et al. An Equation of State for the Thermodynamic Properties of Cyclohexane. **Journal of Physical and Chemical Reference Data**, 43, 2014. 043105.

55. SCHROEDER, J.A.; PENONCELLO, S.G.; SCHROEDER, J.S. A Fundamental Equation of State for Ethanol. **Journal of Physical and Chemical Reference Data**, 43, 2014. 043102.

56. GONÇALVES, A. F.; JUNIOR, E. F. D. C.; DA COSTA, A. O. S. **Comparison of modified alpha functions of the PR-EoS for volume and enthalpy prediction of natural gases**. Annals of the Brazilian Academy of Sciences. Rio de Janeiro: [s.n.]. 2020.

57. KAVIANI, S.; FEYZI, F.; KHOSRAVI, B. Modification of the Peng–Robinson equation of state for helium in low temperature regions. **Physics and Chemistry of Liquids**, 54, December 2015. 1-10.

58. SAALI, A.; SAKHAEINIA, H.; SHOKOUHI, M. Modification of Peng–Robinson Cubic Equation of State with Correction of the Temperature Dependency Term. **Journal of Solution Chemistry**, 50, March 2021. 402-426.

10. Appendix

Table A.1. Group volume (R) and surface area (Q) parameters for each molecular group in the cyclohexane–ethanol system, extracted from DDB database.

Molecular Groups	R	Q
CH_3	0.6325	1.0608
CH_2	0.6325	0.7081
OH	1.2302	0.8927
Cyclic-CH_2	0.7136	0.8635

Table A.2. Group interaction parameters for the UNIFAC model in the cyclohexane - ethanol system, extracted from DDB database (optimization initial guesses).

i	j	a_{ij}	a_{ji}
CH_3/CH_2	OH	986.5	156.4
CH_3/CH_2	Cy-CH_2	-450.4	-34.36
OH	Cy-CH_2	-817.7	1913

Table A.3. Group interaction parameters for the UNIFAC-Do model in the cyclohexane–ethanol system extracted from DDB database.

i	j	a_{ij}	b_{ij}	c_{ij}	a_{ji}	b_{ji}	c_{ji}
CH_3/CH_2	OH	2777	-4.674	0.001551	1606	-4.746	0.0009181
CH_3/CH_2	Cy-CH_2	-117.1	0.5481	-0.00098	170.9	-0.8062	0.001291
OH	Cy-CH_2	3121	-13.69	0.01446	2601	-1.25	-0.006309

Table A.4. Estimated Adjustable Parameters (λ) for the Almeida–Aznar–Telles Cohesion α Function.

Component	λ_1	λ_2	λ_3
Cyclohexane	0.5986	1.0053	0.0799
Ethanol	1.6105	1.079	-0.1457

Table A.5. Estimated UNIFAC Group interaction parameters for the cyclohexane–ethanol system.

i	j	a_{ij}	a_{ji}
CH_3/CH_2	OH	5881.4	6923.4
CH_3/CH_2	Cy-CH_2	-13.3	75.1
OH	Cy-CH_2	344.2	1634.2

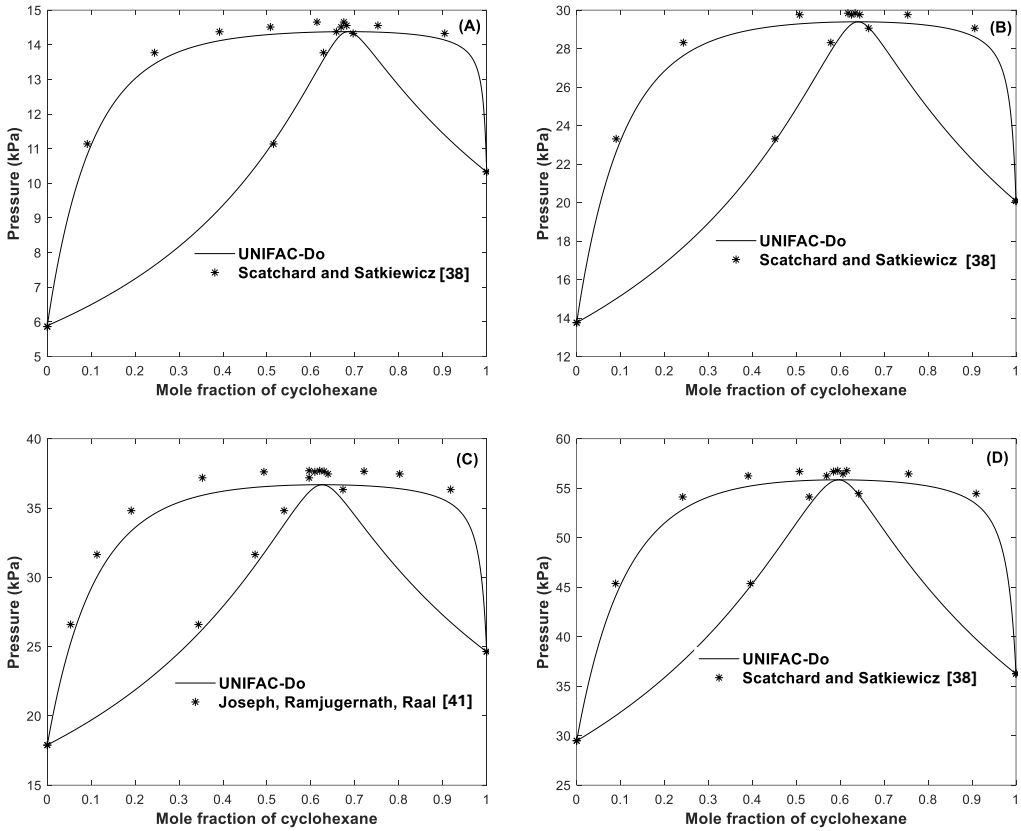

Figure A1. *P*−*z* phase diagrams calculated at 293 K (A), 308 K (B), 313 K (C), and 323 K (D) for the cyclohexane–ethanol system using DDB parameters for UNIFAC-Do.

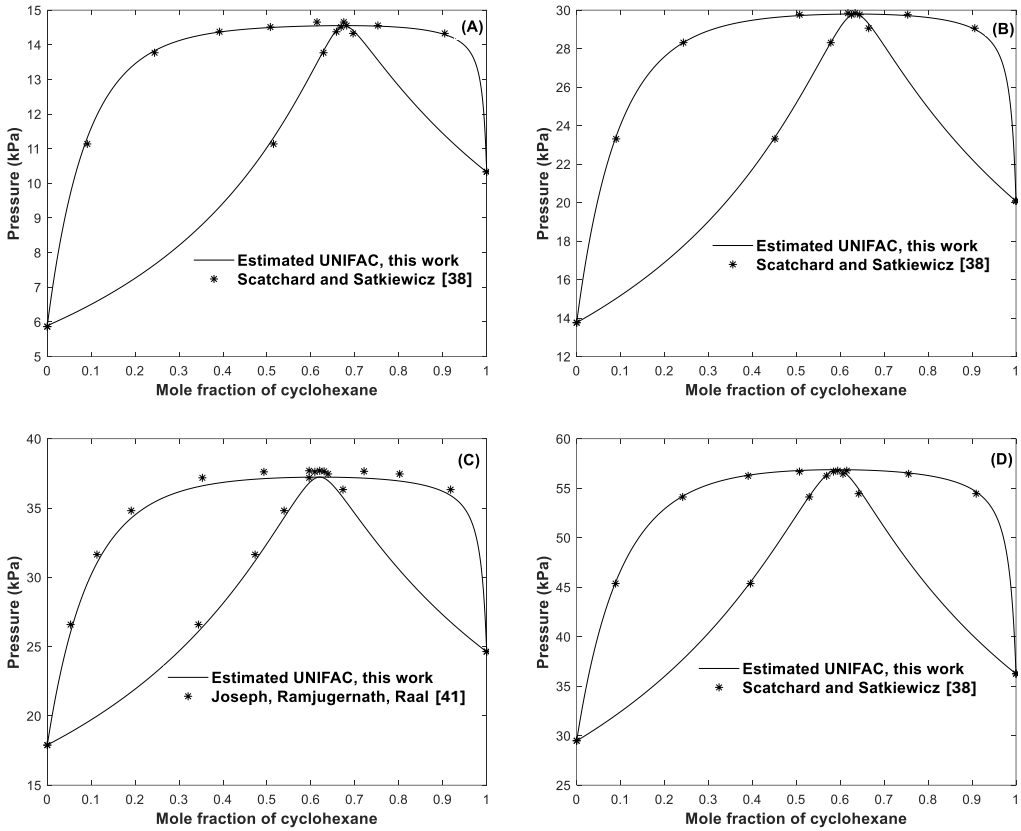

Figure A2. *P*−*z* phase diagrams calculated at 293 K (A), 308 K (B), 313 K (C), and 323 K (D) for the cyclohexane–ethanol system using the group interaction parameters estimated in this work for the UNIFAC model.

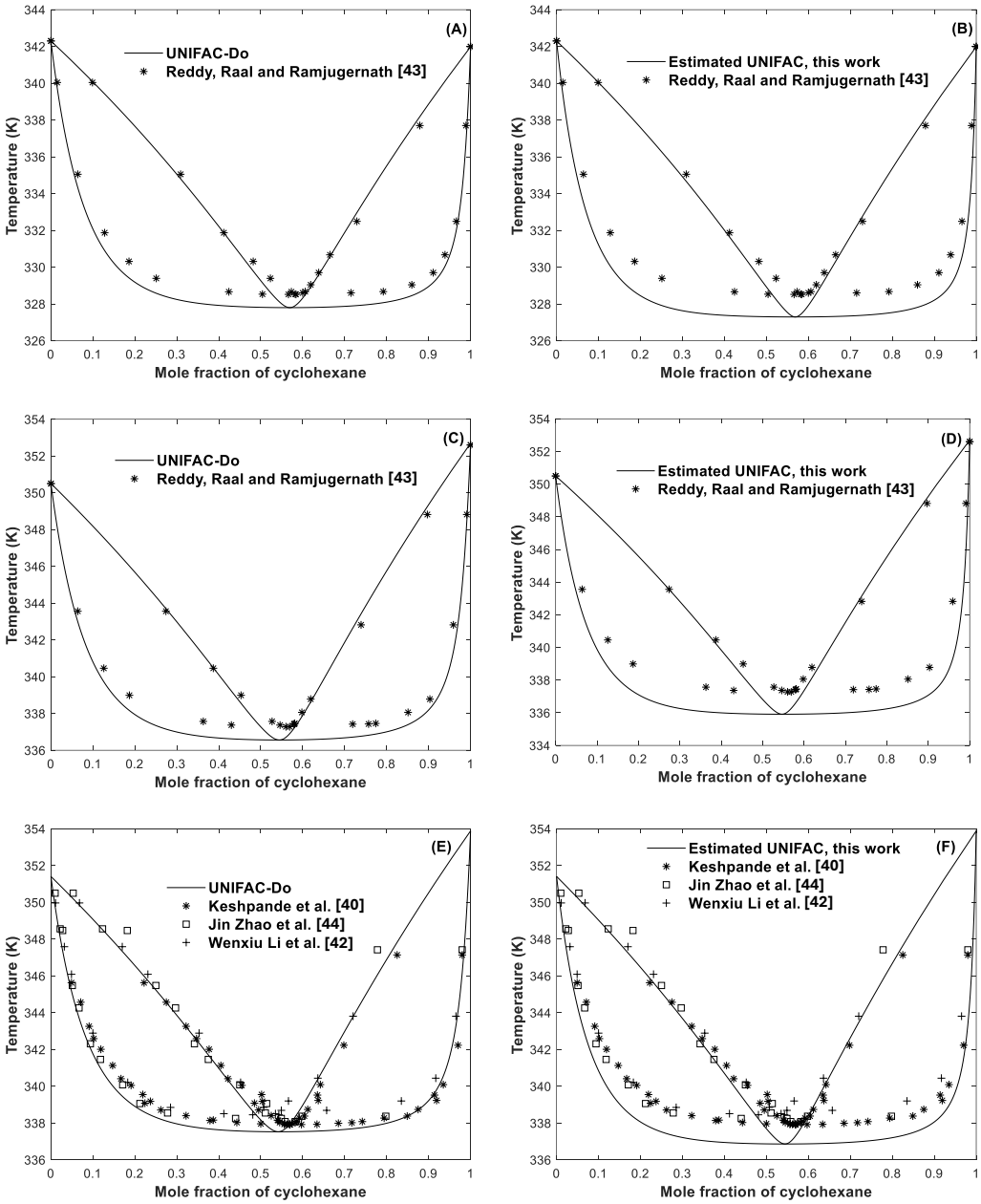

Figure A3. *T*−*z* phase diagrams calculated using DDB UNIFAC-Do and the group interaction parameters estimated in this work for the UNIFAC model at 70 kPa (A and B), 98 kPa (C and D) and 101.3 kPa (E and F).

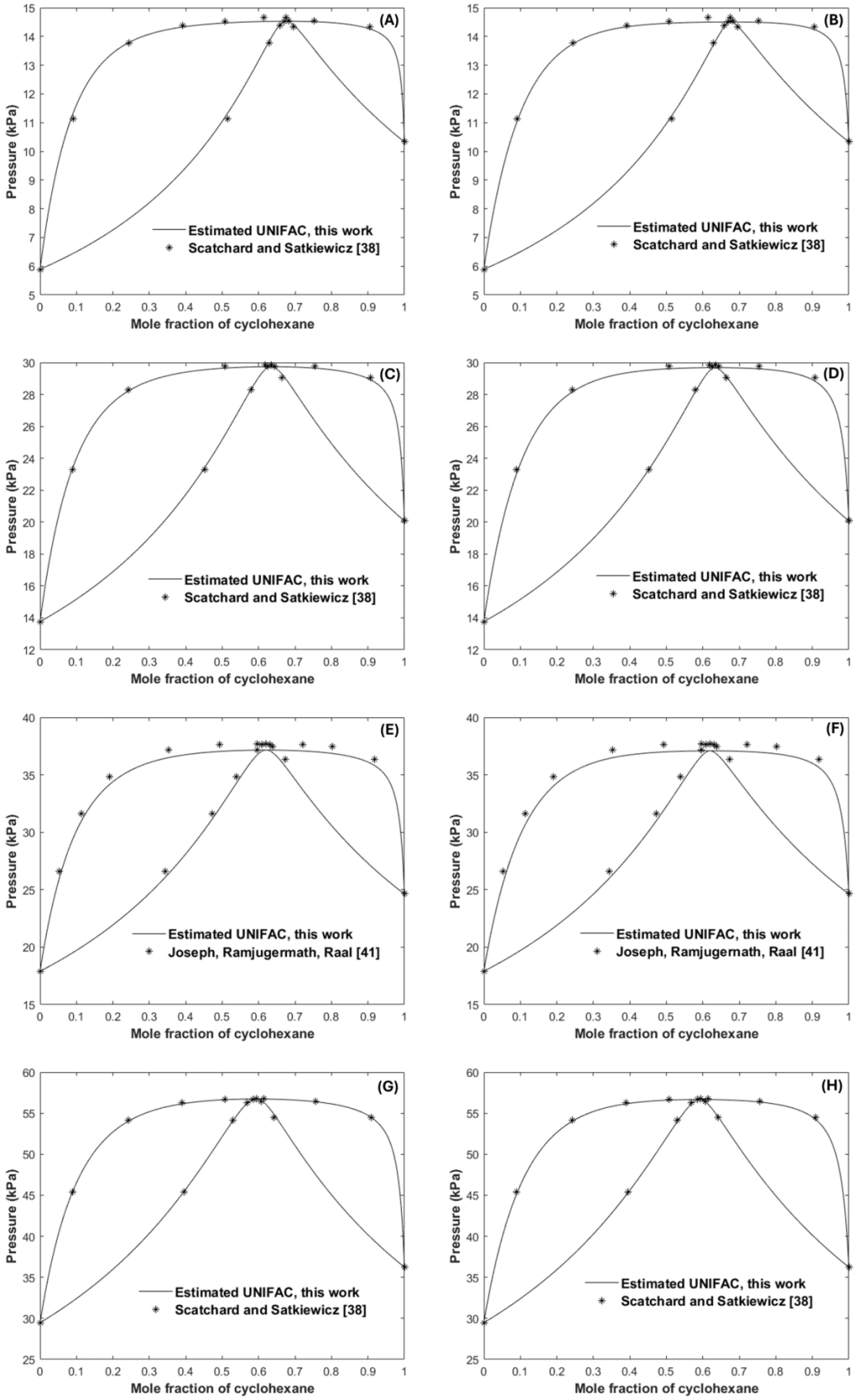

Figure A4. Pressure–composition (*P*–*z*) diagram at 293 K (A and B), 308 K (C and D), 313 K (E and F) and 323 K (G and H) comparing results obtained with group interaction parameters estimated using two different initial guesses: (left) the original DDBST values and (right) the manually modified DDBST values.

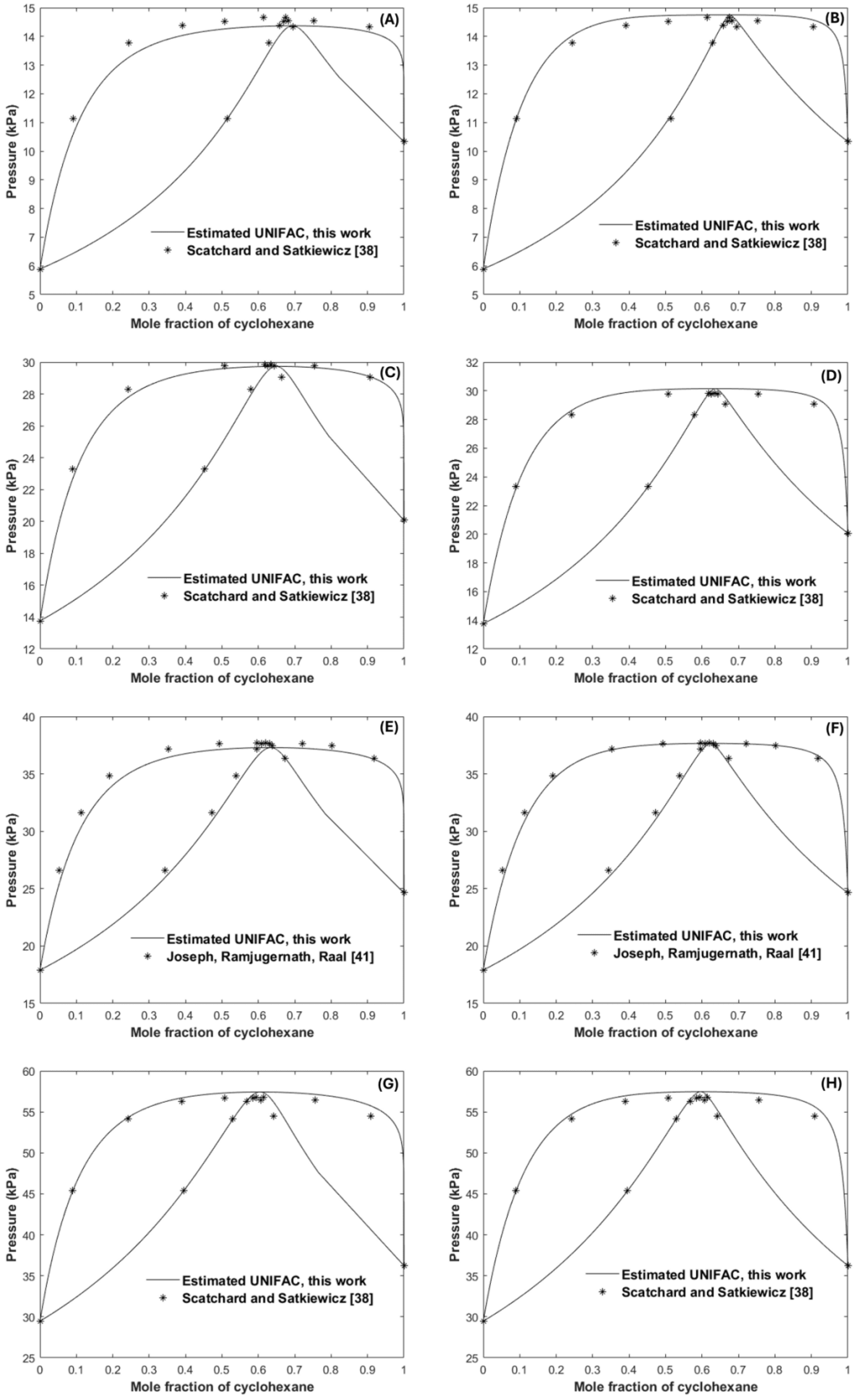

Figure A5. Pressure–composition (*P*–*z*) diagram at 293 K (A and B), 308 K (C and D), 313 K (E and F) and 323 K (G and H) comparing results obtained with group interaction parameters estimated using two different initial guesses: (left) the original DDBST values and (right) the parameter set previously fitted to isothermal VLE data.

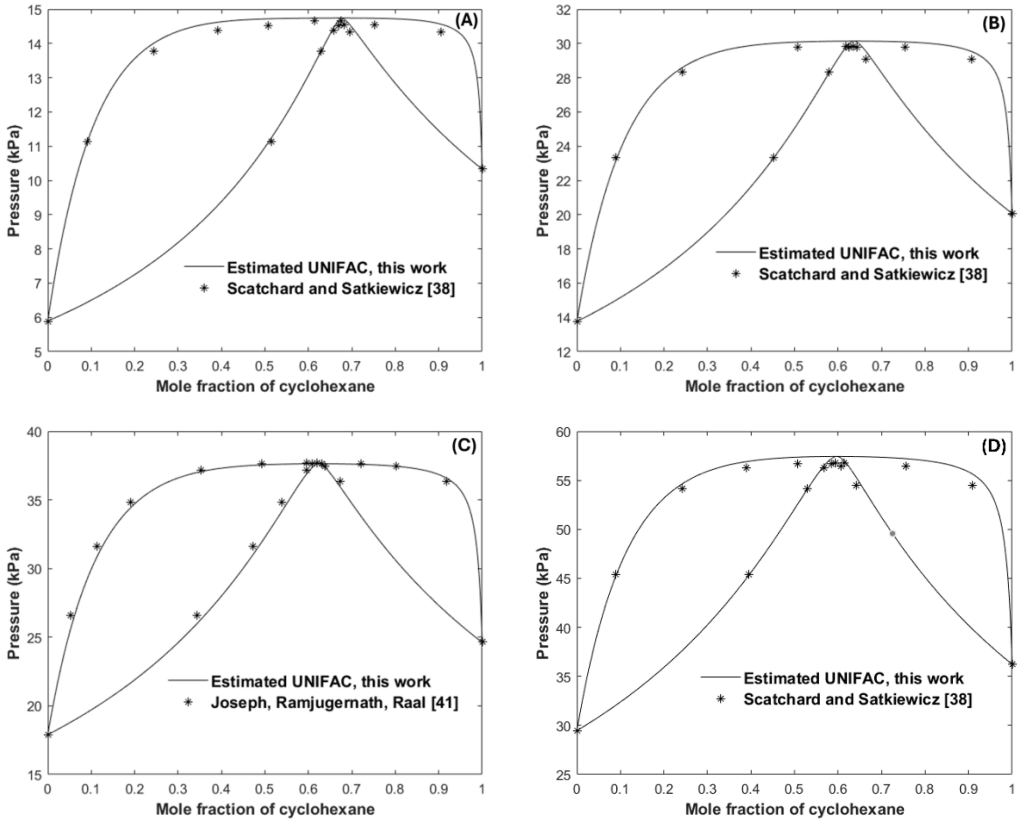

Figure A6. Isothermal *P*–*z* Phase Diagrams for the Cyclohexane–Ethanol Binary System Using UNIFAC Parameters Estimated from COSMO-SAC Data Concentrated in the Cyclohexane-Rich Region: (A) 293 K, (B) 308 K, (C) 313 K, and (D) 323 K.

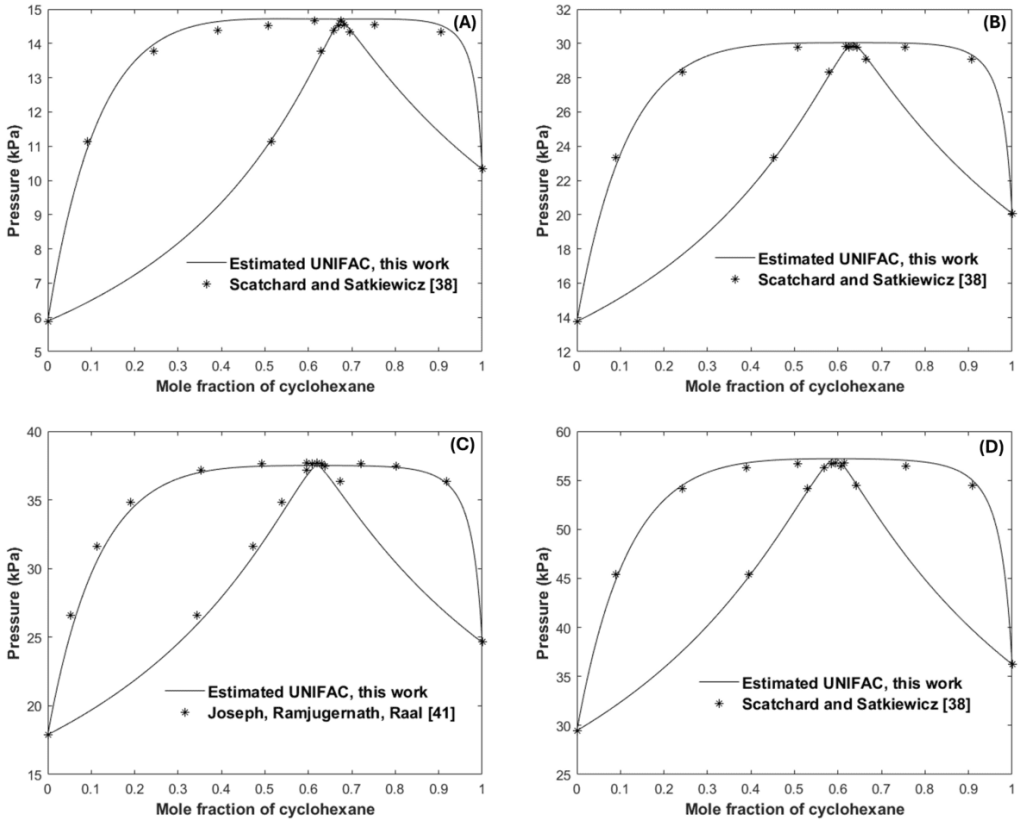

Figure A7. Isothermal *P–z* Phase Diagrams for the Cyclohexane–Ethanol System Using UNIFAC Parameters Optimized with COSMO-SAC Data Concentrated in the Dilute Cyclohexane Region: (A) 293 K, (B) 308 K, (C) 313 K, and (D) 323 K.

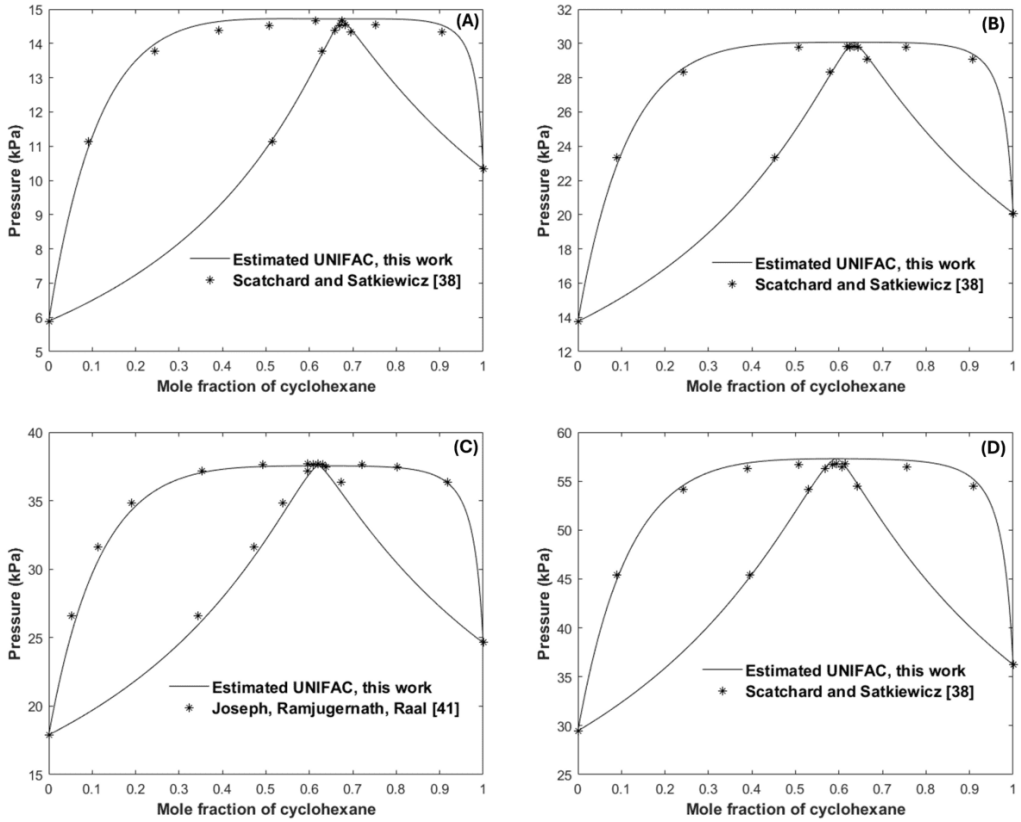

Figure A8. Isothermal *P–z* Phase Diagrams for the Cyclohexane–Ethanol System at (A) 293 K, (B) 308 K, (C) 313 K, and (D) 323 K. Calculations performed using UNIFAC parameters optimized with COSMO-SAC activity coefficients sampled uniformly across the composition range.

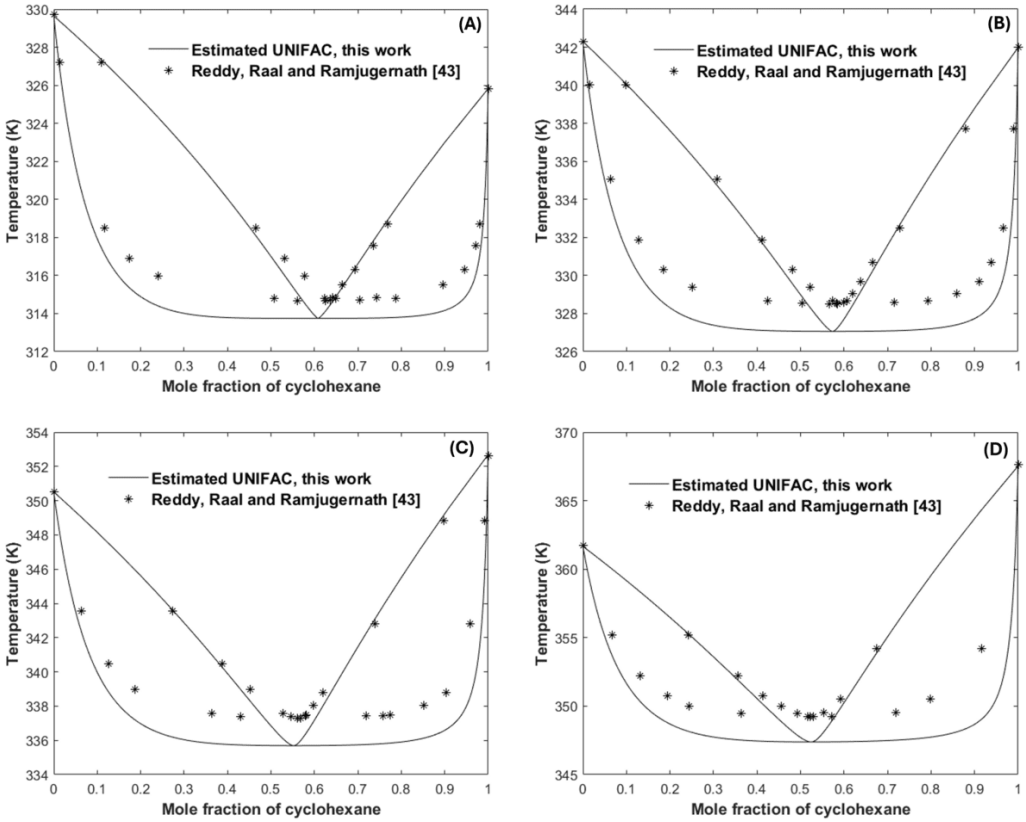

Figure A9. Isobaric *T–z* Phase Diagrams for the Cyclohexane–Ethanol Binary System Using UNIFAC Parameters Estimated from COSMO-SAC Data Concentrated in the Cyclohexane-Rich Region: (A) 40 kPa, (B) 70 kPa, (C) 98 kPa, and (D) 150 kPa.

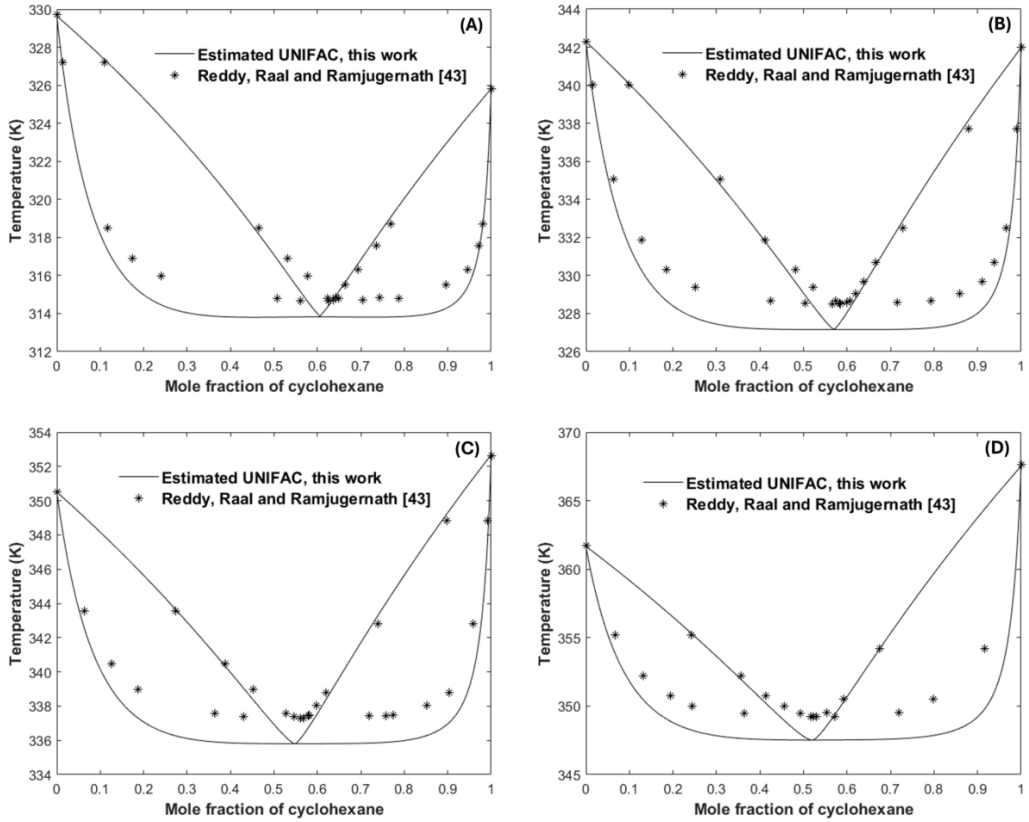

Figure A10. Isobaric *T*–*z* Phase Diagrams for the Cyclohexane–Ethanol System Using UNIFAC Parameters Optimized with COSMO-SAC Data Concentrated in the Dilute Cyclohexane Region: (A) 40 kPa, (B) 70 kPa, (C) 98 kPa, and (D) 150 kPa.

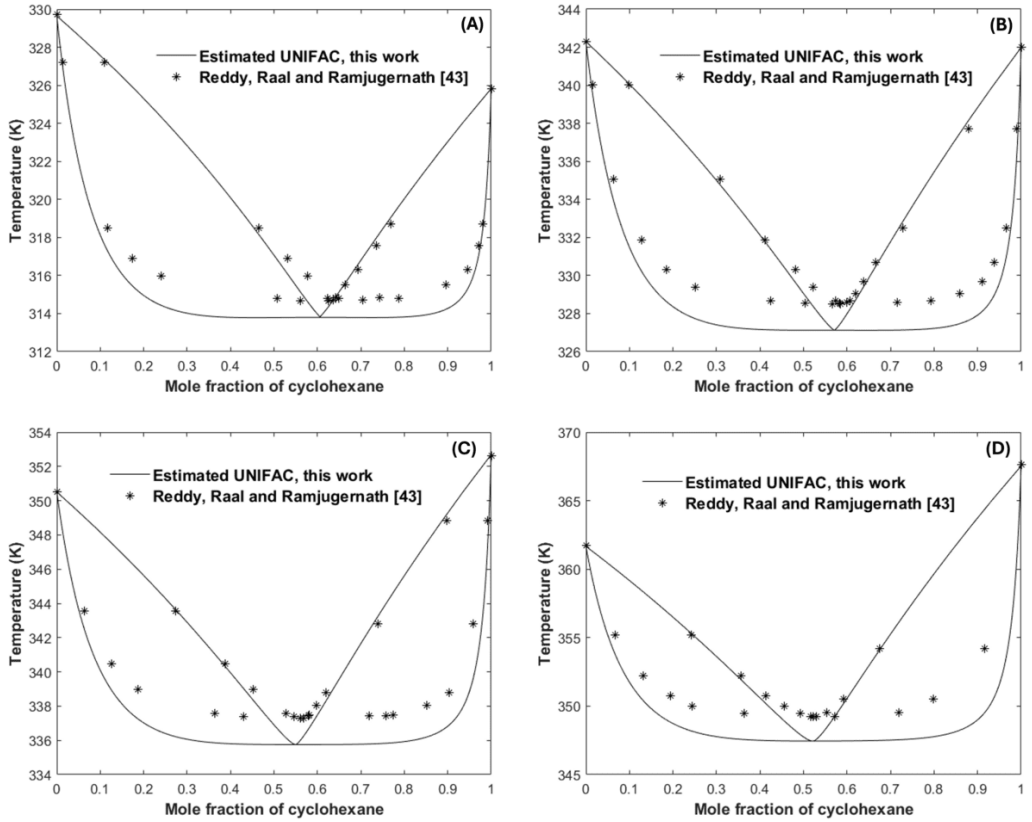

Figure A11. Isobaric *T*–*z* Phase Diagrams for the Cyclohexane–Ethanol System at (A) 40 kPa, (B) 70 kPa, (C) 98 kPa, and (D) 150 kPa. Calculations performed using UNIFAC parameters optimized with COSMO-SAC activity coefficients sampled uniformly across the composition range.