

A systematic global recharacterization method for reservoir fluids in compositional simulations

Barbara F. Esteves^a, Fellipe C. de Oliveira^a, Paulo Couto^b, Paulo L.C. Lage^{a,*}

^a Programa de Engenharia Química — COPPE, Universidade Federal do Rio de Janeiro, 68502, 21941-972, Rio de Janeiro, RJ, Brazil

^b Programa de Engenharia Civil — COPPE, Universidade Federal do Rio de Janeiro, 68506, 21941 909, Rio de Janeiro, RJ, Brazil

ARTICLE INFO

Keywords:

Reservoir compositional simulation
Recharacterization methods
Lumping methods
Complex hydrocarbon mixtures
Quadrature method of moments

ABSTRACT

Component lumping is a common practice to mitigate the computational cost of compositional subsurface simulations. Lumping consists of grouping original components into pseudo-components to reduce the number of components representing a mixture. This work presents a systematic global recharacterization strategy using Lage's method for recharacterizing the C_{6+} or C_{7+} fractions. We compared phase diagram and compositional reservoir simulation results with original and recharacterized fluid compositions. The recharacterized mixtures obtained by the proposed method accurately presented the original mixtures' phase diagram and compositional reservoir behavior. Quantitative error analysis showed that the methodology proposed in this study surpasses a costly global lumping optimization method available in the literature. The replacement of C_+ fractions by only four or three new pseudo-components using Lage's method eliminates the need for lumping important discrete components, such as CO_2 , methane, and ethane, which significantly improved the results. Importantly, the fast methodology presented here offers a means to accurately reduce the number of components characterizing a subsurface fluid from its original composition, independent of the thermodynamic model, and to standardize the mixture recharacterization.

1. Introduction

Compositional reservoir simulations have high computational costs that increase with the number of components used to describe the mixture. Both reservoir simulation and the mixture thermodynamic properties prediction using an equation of state (EoS) play crucial roles (Joergensen and Stenby, 1995; Lolley and Richardson, 1998). The number of mass conservation equations solved in the compositional flow simulation and equilibrium equations necessary for the phase equilibrium calculations is proportional to the number of mixture components. Therefore, the use of a full compositional description of the reservoir fluid is impractical (Joergensen and Stenby, 1995; Lolley and Richardson, 1998; Honami et al., 2000; He and Durlofsky, 2014). Lumping schemes are employed to pseudoize the fluid description (Joergensen and Stenby, 1995; Newley and Merrill, Jr., 1991), substituting its originally identified components by pseudo-components (Joergensen and Stenby, 1995). Consequently, pseudoization may lead to losses in accuracy in the thermodynamic properties prediction and the simulated reservoir flow behavior (Joergensen and Stenby, 1995; Rastegar and Jessen, 2009). Hence, mixture recharacterization methods should accurately describe the phenomena of interest with the smallest number of pseudo-components (Montel and Gouel, 1984).

Lumping strategies for the C_{7+} fraction are readily available in the literature (Pedersen et al., 1985; Danesh et al., 1992; Whitson, 1983) and compared elsewhere (Moghadamzadeh et al., 2013; Oliveira et al., 2023). Particularly, Oliveira et al. (2023) compared three lumping schemes for the recharacterization of the C_{7+} fraction of two condensed gas mixtures assuming the PC-SAFT EoS (Gross and Sadowski, 2001) to the method proposed by Lage (2007). The latter reproduced the phase diagrams constructed with the original mixtures best. Lage's methodology applied the quadrature method of moments (QMOM) to develop the adaptive characterization method, which obtains the fluid pseudo-components using an optimal quadrature rule whose weight function is the mixture molar fraction distribution. That is, Lage's method is extremely accurate in representing fluid properties and adaptive to the fluid compositional description (Lage, 2007; Oliveira et al., 2023).

A standard procedure for global recharacterization of mixtures, however, is yet to be available in the literature. Montel and Gouel (1984) presented a lumping scheme based on the similarities of a few properties of the fluid compounds where an iterative clustering algorithm around mobile centers yields an optimum set of lumps for a given equation of state. Their study used the Peng–Robinson EoS (Peng

* Corresponding author.

E-mail address: paulo@peq.coppe.ufrj.br (P.L.C. Lage).

<https://doi.org/10.1016/j.geoen.2025.213658>

Received 5 June 2024; Received in revised form 12 October 2024; Accepted 2 January 2025

Available online 14 January 2025

2949-8910/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Nomenclature

List of Symbols

c	Factor used in Lage's method
$d_{o, re}$	Euclidean distance between the critical points of original and recharacterized mixtures
e_{β}	Vapor fraction error obtained in a flash condition for the recharacterized mixture.
\bar{e}_{β}	Vapor fraction error for the recharacterized mixture.
e_c	Composition error obtained in a flash condition for the recharacterized mixture.
\bar{e}_c	Composition error for the recharacterized mixture.
$e_{c, phase}$	Composition error obtained in a flash condition for the recharacterized mixture in the liquid or vapor phase.
I_i	Dimensionless molar mass
M	Molar mass
M^0	Molar mass of components in the original mixture
N	Total number of components in the mixture
N^0	Number of components in the original mixture
N_d	Number of all components but the plus fraction in the original mixture
n_l	Number of components lighter than the C_+ fraction
w	Weight of the quadrature
x	Molar fraction
z	Normalized molar fraction

Abbreviations

CCS	CO ₂ capture and storage
SCN	Single carbon number

Greek letters

β	Vapor fraction
μ	Moment of the normalized discrete distribution
Θ	Property that depends primarily on the molar mass
ξ	Abscissa of the quadrature

Subscripts

c	Critical property
o	Original fluid
rec	Recharacterized fluid

and Robinson, 1976), but the principles of the method do not depend on the selected EoS. They compared some results obtained with the mixture characterization from the proposed methodology with those using the original fluid composition and a classical lumping method based on experience, which includes tuning to experimental data. The results were for the formation volume factor and the amount of dissolved gas in a differential liberation experiment (for the oil), the liquid deposit in a constant mass depletion experiment, and the global gas compressibility factor. Their results, considering 5 and 7 pseudo-components, show the poor accuracy of the lumping scheme based on experience.

Newley and Merrill, Jr. (1991) developed a lumping scheme based on minimizing the difference between the equilibrium ratios of the original components and the pseudo-component to which they are assigned. For this, they ordered the components by equilibrium ratio values and assumed an initial component grouping determined on an equal-mole-fraction basis. They obtained good results for the lumped mixture models compared to the original mixture composition through PVT data, phase behavior curves, and one-dimensional compositional simulation results, such as reservoir pressure, oil recovery, gas-oil ratio, and liquid saturation.

Joergensen and Stenby (1995) compared three different grouping methods to formulate a generalized set of rules for the pseudoization of reservoir fluids. The three selected grouping methods were: (A) equality of mole fractions, where the original components of the fluid are ordered concerning their normal boiling point and grouped to form pseudo-components with approximately equal mole fractions (Kay, 1936); (B) equality of weight fractions, similar to (A) but forming pseudo-components with equal weight fractions (Pedersen et al., 1985); and (C) multi-variant lumping, which groups pure components with similar properties using an iterative scheme, in which the distances between the pure components and the pseudo-components properties are minimized (Montel and Gouel, 1984). Joergensen and Stenby (1995) employed the SRK EoS and evaluated the lumping schemes by simulating phase equilibrium and PVT experiments. Two grouping rules were assumed: never group methane and do not group ethane with components heavier than propane. The accuracy of the analyzed lumping schemes depended primarily on the number of pseudo-components used in the fluid descriptions. For their experimental data, they showed that the optimal number of pseudo-components ranged from six to eight.

Mello et al. (2011) evaluated the impact of simplifications in the fluid characterization on reservoir compositional simulations, particularly the effects of oil production predictions and computational costs. They proposed a variety of lumping schemes and analyzed two different tuning techniques (Coats and Smart, 1986; Pedersen et al., 1988), each one considering a different EoS for the fluid characterization. They employed a reservoir model based on the SPE 10th (Christie and Blunt, 2001) comparative study and the oil data from Pedersen et al. (1989). They concluded that the number of pseudo-components, the lumping scheme, and the tuning method are all crucial for correctly reproducing oil's original data.

Alavian et al. (2014) presented a global pseudoization procedure based on (i) selecting all plausible lumping combinations assuming a few imposed constraints and (ii) quantifying the prediction errors of some experimental data using the chosen EoS and the lumped fluid composition. Their method proposes finding an optimal pseudoized EoS model to a particular reservoir development. The desired number of pseudo-components, the thermodynamic model, and the experimental data are inputs of the procedure. They evaluate the accuracy of the lumped model by a weighted root mean square (RMS) difference between experimental and calculated PVT properties, employing the Peng–Robinson EoS and Lorentz–Bray–Clark viscosity correlation (Lohrenz et al., 1964). They employed the PhazeComp EOS Modeling software (Zick Technologies, 2024) correlations for calculating pseudo-components properties. By analyzing five reservoir fluids, Alavian et al. (2014) concluded that their methodology was successful in generating lumped fluid compositions that reproduce well the behavior of the original fluid model.

Some suggestions exist on how to evaluate the performance of lumping schemes. Lolley and Richardson (1998) proposed a guideline for accepting a pseudoized mixture. They assumed that the recharacterized fluid had to match PVT and laboratory data within 5% and must give a simulated reservoir production behavior that agrees within 10% to those using the original fluid. This study used a commercial phase calculation program to tune the EoS parameters and pseudoization calculations. They assumed the Peng–Robinson EoS and used vapor

pressure, steam distillation, and viscosity measurements to tune the EoS parameters. The base case was a forty-component fluid characterization of a heavy oil mixture. The results showed that a minimum of four pseudo-components were required to simulate thermal processes and recovery mechanisms accurately in the steam flooding of this heavy oil based on the criteria defined in the study. Honami et al. (2000) reported results of accuracy and efficiency of lumping schemes regarding compositional reservoir simulation for datasets of gas condensates and volatile oils. They observed that the number of pseudo-components did not affect the predictions of pressure, composition, and production performance, which were affected by the quality of PVT data regression. However, the CPU time of flash calculations increased exponentially with the number of pseudo-components.

Although, several lumping schemes have been defined and investigated, particularly using cubic EoSs and critical properties correlations, a standard and accurate methodology decoupled from tuning processes and depending only on composition data would be particularly interesting. Therefore, this study proposes a systematic and accurate methodology to recharacterize the original fluid composition to reduce the number of components describing the fluid phase, consequently reducing simulation times, and making compositional reservoir simulations faster, viable, and more trustful.

For this, here we propose lumping strategies for the discrete components based on literature and research insights and the use of the recharacterization method proposed by Lage (2007) for the C_{6+} or C_{7+} fractions. Importantly, this is the first time the method of Lage (2007) is used for compositional reservoir simulations. We employed the Peng–Robinson EoS for PVT calculations, generating phase diagrams with the original and recharacterized mixture compositions. The results are discussed based on quantitative error analysis and comparisons to literature data. Finally, we conducted compositional reservoir simulations, comparing results for gas–oil ratio, hydrocarbon recovery, and CO_2 composition.

2. Methodology

This section describes the global recharacterization method proposed in this study and the methodology developed by Alavian et al. (2014) used for comparison. It also introduces the PVT EoS, mixtures' characterization, reservoir models, and numerical procedures.

2.1. Global recharacterization method

The proposed global lumping scheme for the mixture composition combines some grouping rules for the discrete components with the recharacterization method of Lage (2007) for the C_{7+} or C_{6+} fractions. Importantly, this method does not include any re-tuning of the EoS parameters for the new pseudo-components. Therefore, it is independent of the thermodynamic model and requires no experimental data.

2.1.1. Single carbon number (SCN) components

The first step of our methodology is lumping chemical components with six or more carbons into single carbon number (SCN) components (C_6 , C_7 , and others) if such non-lumped components are present in the original mixture composition.

2.1.2. Recharacterization method of Lage (2007) for the C_+ fraction

Consider the fluid composition with N^0 components with molar mass M_i^0 and molar fractions x_i^0 , $i = 1, \dots, N^0$, in which the plus fraction (C_{6+} or C_{7+}) has been already described using SCN components. The last component ($i = N^0$) is usually a heavy plus fraction (C_{20+} , C_{30+} , or even heavier).

The method recharacterizes the plus fraction (C_{6+} or C_{7+}) described by SCN components using the adaptive recharacterization method developed by Lage (2007). Its main assumption is that the thermodynamic model parameters of the plus fraction are primarily a function of the

molar mass, for which correlations are available.

The SCN components of the C_{6+} or C_{7+} fractions belong to the index range $[N_d + 1, N^0]$, where N_d is the number of original mixture discrete components that will form n_l lumps using the grouping rules described in the following. Thus, the recharacterized mixture has $N = n_l + n_p$ components with molar mass M_i , and molar fraction x_i , $i = 1, \dots, N$, being n_p the chosen number of new pseudo-components used to recharacterize the plus fraction (Oliveira et al., 2023). Usually, $n_p = 2, 3$ or 4 suffices.

The method of Lage (2007) calculates the Gauss–Christoffel quadrature rule from the moments of the molar fraction discrete distribution in terms of the molar mass. The Gauss–Christoffel quadrature is a Gaussian quadrature whose weighting function is an unknown distribution, for which only their moments are known. If $2N$ of such moments are available, one can calculate the N -point Gauss–Christoffel quadrature using several methods (Chicralla et al., 2019; John and Thein, 2012). We can interpret the resulting N -point Gauss–Christoffel quadrature as a new discrete distribution defined by the weighted sum of N Dirac delta distributions at the quadrature abscissas using the corresponding quadrature weights. This new discrete distribution has the same values for the first $2N$ moments of the original distribution. Thus, it is a discrete distribution that approximates the original distribution.

For calculating the Gauss–Christoffel quadrature, it is convenient to normalize the molar fractions of the plus fraction:

$$z_i = \frac{x_i}{\sum_{j=N_d+1}^{N^0} x_j}, \quad (1)$$

and dimensionless molar masses defined by

$$I_i = c \frac{M_i^0 - M_{N_d+1}^0}{M_{N^0}^0 - M_{N_d+1}^0} \quad (2)$$

where c is a factor close to one chosen to make the required moments of the discrete distribution $\{z_i, I_i\}_{i=N_d+1}^{N^0}$ to have values close to one (Chicralla et al., 2019), which are given by:

$$\mu_k = \sum_{i=N_d+1}^{N^0} z_i I_i^k, \quad \forall k \quad (3)$$

One can compute the n_p -point Gauss–Christoffel quadrature rule from the $2n_p$ first moments (μ_k , $k = 0, 1, \dots, 2n_p - 1$) by several methods, which are reviewed elsewhere (Chicralla et al., 2019; John and Thein, 2012). Let ξ_j and w_j , $j = 1, \dots, n_p$, be the abscissas and weights of this quadrature. Then, the new pseudo-components that represent the plus fraction in the global recharacterized mixture have the molar masses given by:

$$M_{n_l+i} = \xi_i \frac{M_{N^0}^0 - M_{N_d+1}^0}{c} + M_{N_d+1}^0, \quad i = 1, \dots, n_p \quad (4)$$

whose molar fractions are

$$x_{n_l+i} = w_i \sum_{j=N_d+1}^{N^0} x_j^0, \quad i = 1, \dots, n_p \quad (5)$$

For any property that depends primarily on the molar mass, $\theta(M)$, the corresponding C_+ fraction property is given by:

$$\Theta = \sum_{j=N_d+1}^{N^0} \theta(M_j^0) x_j^0 \quad (6)$$

that can be accurately approximated by

$$\Theta \cong \sum_{j=1}^{n_p} x_{n_l+j} \theta(M_{n_l+j}) \quad (7)$$

using the recharacterized mixture composition.

The reader can find more details in the works of Lage (2007), Chicralla et al. (2019), and Oliveira et al. (2023). Additionally, the code for implementing Lage's method is freely available in a public repository (Lage, 2024).

2.1.3. Grouping rules for remaining discrete components

We defined four grouping rules for the proposed global recharacterization method using the literature and insights obtained throughout this investigation. Using $n_p = 3$ or 4 for the C_+ fraction, these rules gave recharacterized mixtures with 8 or 9 components that accurately represented the PVT properties of the analyzed mixtures and were amenable to performing reservoir compositional simulations.

1. Carbon dioxide (CO_2) remains ungrouped due to the imperative need for accurate and reliable CO_2 thermodynamic properties for the development of CO_2 capture and storage (CCS) technologies (Wang et al., 2019). This approach was also assumed in other studies (Montel and Gouel, 1984; Alavian et al., 2014; Mello et al., 2011).
2. Methane (CH_4) groups only with nitrogen (N_2) because methane plays a crucial role in understanding the phase behavior of most reservoir fluids (Joergensen and Stenby, 1995) and N_2 is usually present in trace amounts.
3. Ethane (C_2H_6) and propane (C_3H_8) remain ungrouped. Although Joergensen and Stenby (1995) suggests grouping C_2H_6 and C_3H_8 , we verified the importance of keeping these components ungrouped in the recharacterized mixture.
4. Components with four and five carbon atoms form a single group when applying the method of Lage (2007) to the C_{6+} fraction. Similarly, components with four, five, and six carbons form a single group when recharacterizing the C_{7+} fraction.

If a mixture with fewer components is necessary, one should first consider lumping C_2H_6 and C_3H_8 .

2.2. Global lumping method of Alavian et al. (2014)

In order to compare our methodology with another method, we selected the results from the global component lumping method developed by Alavian et al. (2014). This method uses optimization techniques to find the best match between the original and pseudoized mixtures' results among all feasible lumping combinations. Besides, the best lumping scheme also depends on the weight factors used in the optimization objective function to weigh the errors of the different mixture properties. The reader should consult the original reference for more details on this method.

2.3. PVT simulation

We calculated phase diagrams to compare fluid behaviors obtained by the recharacterized and original fluid compositions using the commercial PVT simulator WinProp/CMG (Computer Modelling Group Ltd., 2022). After defining the fluid components and their molar fractions, we chose the EoS model and correlations to determine fluid properties. Only cubic EoSs are available in this software, which usually occurs for most commercial reservoir simulators.

2.3.1. Peng–Robinson EoS

We selected the Peng and Robinson (1976) EoS, a commonly used thermodynamic model. The binary interaction parameters (BIP) were assumed null, fluid viscosities came from Pedersen viscosity correlation (Pedersen and Fredenslund, 1987), and the critical properties and acentric factor for pure components were calculated using Twu (1984) and Lee and Kesler (1975) correlations, respectively.

We calculated the critical properties for discrete components or pseudo-components as described below.

- Pure components' critical properties came directly from the literature.
- Original SCN components' critical properties were obtained using correlations based on the molar mass (Hosseinfar and Jamshidi, 2014). See Appendix A for more details.

Table 1

Flash points selected for error analysis.

Flash point	T (°C)	P (kPa)	Description
P1	50	10 000	Close to the bubble point curve
P2	200	10 000	Far from the phase envelope
P3	250	18 000	Close to bubble point curve
P4	363	14 042	0.95 P_c and 0.95 T_c
P5	450	80 000	Close to the dew point curve

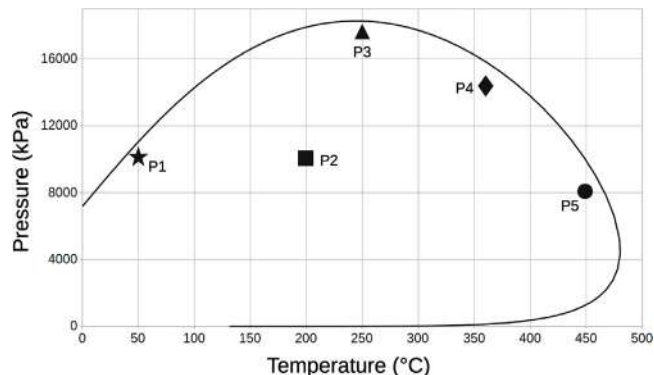


Fig. 1. Phase diagram for the original composition of Mixture B and location of the five flash points used in the error analysis.

- Lumped components' critical properties came from the molar averaging of the original components' properties (Kay, 1936).
- Re-characterized components' critical properties also came from correlations based on the molar mass (Hosseinfar and Jamshidi, 2014).

We assumed the simple molar averaging strategy for obtaining molar masses and critical properties of lumped components based on Joergensen and Stenby (1995), who tested four property calculation methods for lumped components derived from pure components: molar averaging, weight-based average, and mixing-rule based calculations with and without temperature dependency. They showed that the method based on mixing rules without temperature dependency performed worse than the others.

2.4. Error analysis

In order to quantify the accuracy of the recharacterized compositions obtained using our global recharacterization method and the method developed by Alavian et al. (2014), we performed an error analysis for Mixture B considering the accuracy of the critical points (P_c and T_c), and the errors in vapor fraction (β) and compositions in flash separations.

Table 1 gives the five selected flash conditions, and Fig. 1 shows their location within the phase diagram of the original mixture.

2.4.1. Critical point error

The error between the predicted values of P_c and T_c using the original and recharacterized mixture compositions is defined by the Euclidean distance ($d_{o,rec}$) between these points in the normalized plane

$$\frac{P}{P_c} \times \frac{T}{T_c} \text{ given by: } d_{o,rec} = \sqrt{\left(1 - \frac{P_{c,rec}}{P_{c,o}}\right)^2 + \left(1 - \frac{T_{c,rec}}{T_{c,o}}\right)^2} \quad (8)$$

where $P_{c,rec}$ and $T_{c,rec}$ are the critical pressure and temperature of the recharacterized mixture and $P_{c,o}$ and $T_{c,o}$ represent these properties for the original mixture.

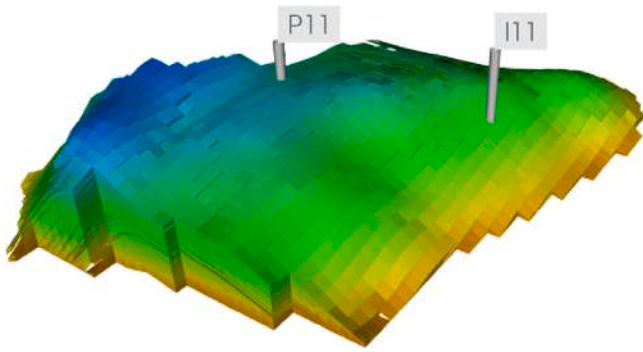


Fig. 2. Reservoir grid top view locating the injector (I11) and production (P11) wells. The reservoir is shown with a vertical exaggeration of 3x.

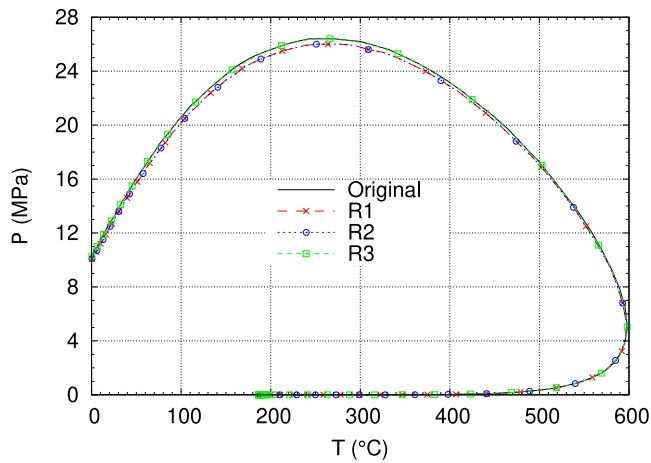


Fig. 3. Phase diagrams for Mixture A using the original mixture composition and recharacterizations in which the light components (C_2 to C_5) are grouped or not.

2.4.2. Vapor fraction error (\bar{e}_β)

We calculated the absolute error in β prediction for each flash condition (e_β) by comparing the results for the original and recharacterized mixtures. Then, we determined the mean vapor fraction error (\bar{e}_β) as the arithmetic average of the e_β values for the five flash conditions selected for this study.

2.4.3. Composition error (\bar{e}_c)

To compare the compositional error between the results of a mixture recharacterization and its original composition, we summed the molar compositions obtained for the original mixture corresponding to each recharacterized pseudo-component. For instance, if one recharacterization lumps C_2 and C_3 , the lumped molar fraction is compared to the sum of the molar fractions of C_2 and C_3 in the original mixture. For the C_+ fraction, we summed the molar fractions of all pseudo-components in this fraction in both original and recharacterized mixtures and then compared them.

After determining the errors in the new pseudo-components molar fractions, we calculated the arithmetic average of their absolute values for each phase to determine the phase compositional error ($e_{c,phase}$), whose arithmetic average defines the compositional error of a flash calculation (e_c). The overall composition error \bar{e}_c is the average of e_c for the selected five flash conditions.

2.5. Reservoir simulator

This study employed the commercial compositional reservoir simulator GEM/CMG (Computer Modelling Group Ltd., 2022). For reference, when discussing run-time in the results section, a computer with

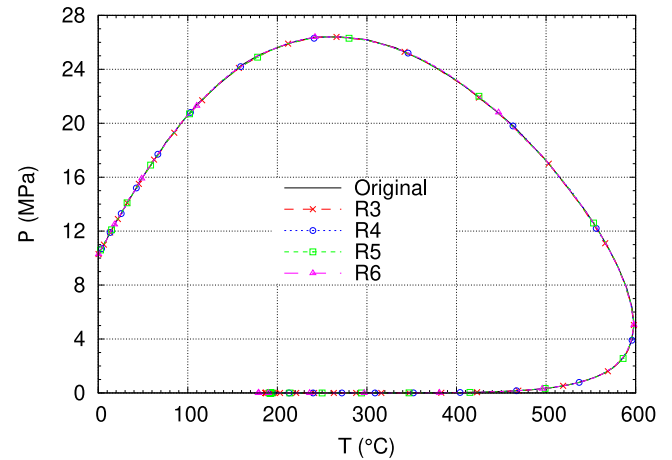


Fig. 4. Phase diagrams for Mixture A using the original mixture composition and recharacterizations that the C_{6+} or C_{7+} fractions are represented using three or four pseudo-components.

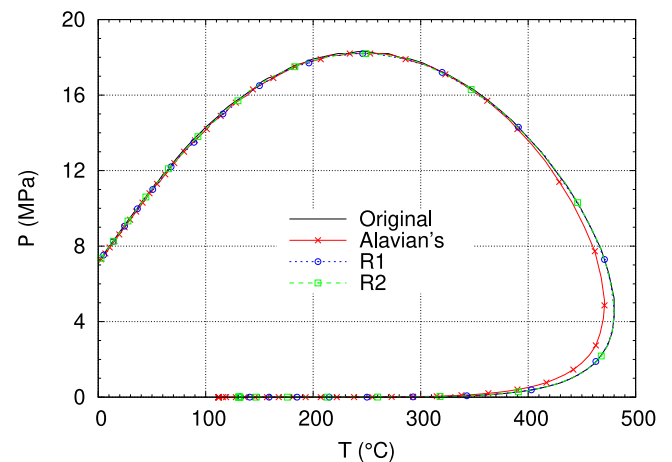


Fig. 5. Phase diagrams for Mixture B using the original mixture composition and recharacterizations given by our method and that proposed by Alavian et al. (2014).

an Intel(R) Core(TM) i7-8750H CPU @ 2.20 GHz was used to perform the simulations and an average value of three runs is presented. In all simulations, we choose the single-point upwinding scheme, which is first-order accurate, for spatial discretization, and an adaptive implicit method that blends forward and backward Euler schemes, which are both first-order accurate, for time discretization.

2.6. Mixtures and reservoir model

This study considered two mixtures whose original compositions are given below. The reservoir model used to perform compositional simulations is also presented in this section.

2.6.1. Mixture A

Mixture A was a recombined live oil that original composition came from the gas chromatography of its flash products (Costa et al., 2022). Appendix B describes Mixture A, including its components, their molar masses, molar fractions, and critical properties.

We compared six recharacterization schemes for this fluid in order to evaluate the effects on its phase equilibrium behavior when assuming: (a) C_2 and C_3 were either grouped or ungrouped in the recharacterized mixture, (b) C_4 and C_5 were either grouped or ungrouped in the recharacterized mixture, and (c) using three or four

Table 2
Components of Mixture A.

Characterization	N	Components
Original	41	CO ₂ , CH ₄ , C ₂ H ₆ , C ₃ H ₈ , iC ₄ , nC ₄ , iC ₅ , nC ₅ , C ₆ , m-cyclo-C ₅ , benzene, cyclo-C ₆ , C ₇ , m-cyclo-C ₆ , toluene, C ₈ , C ₂ -benzene, m&p-xylene, o-xylene, C ₉ , C ₁₀ , C ₁₁ , C ₁₂ , C ₁₃ , C ₁₄ , C ₁₅ , C ₁₆ , C ₁₇ , C ₁₈ , C ₁₉ , C ₂₀ , C ₂₁ , C ₂₂ , C ₂₃ , C ₂₄ , C ₂₅ , C ₂₆ , C ₂₇ , C ₂₈ , C ₂₉ , C ₃₀₊
R1	7	CO ₂ , C ₁ , C ₂ - C ₃ , C ₄ - C ₅ , and 3 pseudos for C ₆₊
R2	10	CO ₂ , C ₁ , C ₂ - C ₃ , iC ₄ , nC ₄ , iC ₅ , nC ₅ , and 3 pseudos for C ₆₊
R3	8	CO ₂ , C ₁ , C ₂ , C ₃ , C ₄ - C ₅ , and 3 pseudos for C ₆₊
R4	9	CO ₂ , C ₁ , C ₂ , C ₃ , C ₄ - C ₅ , and 4 pseudos for C ₆₊
R5	8	CO ₂ , C ₁ , C ₂ , C ₃ , C ₄ - C ₆ , and 3 pseudos for C ₇₊
R6	9	CO ₂ , C ₁ , C ₂ , C ₃ , C ₄ - C ₆ , and 4 pseudos for C ₇₊

Table 3
Components of Mixture B.

Characterization	N	Components
Original	34	N ₂ , CO ₂ , CH ₄ , C ₂ H ₆ , C ₃ H ₈ , iC ₄ , nC ₄ , iC ₅ , nC ₅ , C ₆ , C ₇ , C ₈ , C ₉ , C ₁₀ , C ₁₁ , C ₁₂ , C ₁₃ , C ₁₄ , C ₁₅ , C ₁₆ , C ₁₇ , C ₁₈ , C ₁₉ , C ₂₀ , C ₂₁ , C ₂₂ , C ₂₃ , C ₂₄ , C ₂₅ , C ₂₆ , C ₂₇ , C ₂₈ , C ₂₉ , C ₃₀₊
Alavian's	9	N ₂ , CO ₂ , C ₁ , C ₂ - C ₃ , C ₄ - C ₆ , C ₇ - C ₁₁ , C ₁₂ - C ₁₅ , C ₁₆ - C ₂₃ , C ₂₄ - C ₃₀₊
R1	9	CO ₂ , N ₂ - C ₁ , C ₂ , C ₃ , C ₄ - C ₆ , and 4 pseudos for C ₇₊
R2	9	CO ₂ , N ₂ - C ₁ , C ₂ , C ₃ , C ₄ - C ₅ , and 4 pseudos for C ₆₊

pseudo-components representing the C₊ fraction. Table 2 summarizes the analyzed characterizations for Mixture A, where N is the total number of components in the mixture.

2.6.2. Mixture B

Mixture B is a reservoir fluid generated from a gas condensate composition using an isothermal chemical-gravity segregation model, described in the work of Alavian et al. (2014) as their less-volatile oil (LVO) with solution GOR of 1000 scf/STB. Appendix C gives Mixture B's complete composition and thermodynamic properties of its thirty-four components. The optimization strategy of Alavian et al. (2014) generated the best lumping scheme for fluid B with nine components. The goal of using Mixture B was to compare the proposed global recharacterization strategy to the one proposed by Alavian et al. (2014) using two recharacterization schemes with nine components, as in the best-lumped mixture given by Alavian et al. (2014). All mixture B's characterizations are briefly described in Table 3.

2.6.3. Reservoir model

The benchmark model UNISIM-IV-2019 (Botchia et al., 2023) for subsurface reservoirs was used in this study, with some simplifications. The model has a grid resolution of 200 × 200 × 5 m, with 47 × 39 × 291 blocks in 'i', 'j' and 'k' directions, respectively, totaling 533,403 blocks, from which 63,694 are active in our simulations. Fig. 2 shows the reservoir model depicting the locations of the injector and producer wells. A maximum surface gas rate constraint of 4,000,000 m³/day for the injector well and a maximum total surface liquid rate constraint of 8000 m³/day for the producer well were assumed. Porosity and permeability assumed a normal distribution and a log-normal distribution, respectively. Other simulation conditions are: the rock compressibility at 50,000 kPa was 5 × 10⁻⁷ kPa⁻¹, the reservoir pressure at 5300 m was 62,000 kPa, and the reservoir temperature was 100 °C. The reader can find more details in the work of Botchia et al. (2023).

3. Results and discussion

This section compares the results for the phase diagrams and compositional reservoir simulations obtained using the original mixtures' compositions and their recharacterizations.

3.1. Phase diagrams

In order to determine the best lumping scheme for the discrete components, firstly, we compared phase diagrams aiming to evaluate the reproducibility of using recharacterized compositions for Mixture A. Fig. 3 presents the phase diagrams constructed considering the original composition of Mixture A, and recharacterizations R1, R2, and R3 described in Table 2.

Fig. 3 clarifies the relevance of maintaining ethane (C₂H₆) and propane (C₃H₈) ungrouped in the mixture recharacterization, as can be seen by the red curve (R3). Furthermore, the accuracy obtained by recharacterization R3, with only eight components, effectively reproducing the behavior of the original fluid composition with 41 components, was only possible because the highly accurate method developed by Lage (2007) used to recharacterize the C₆₊ fraction with only three new pseudo-components.

Fig. 4 demonstrates that the global recharacterization strategy works properly using the method of Lage (2007) for the C₆₊ or C₇₊ fractions recharacterizations using three and four pseudo-components. As can be seen, global recharacterizations R3 to R6, ranging between eight and nine components, reproduced the behavior of the original composition of Mixture A (41 components) well.

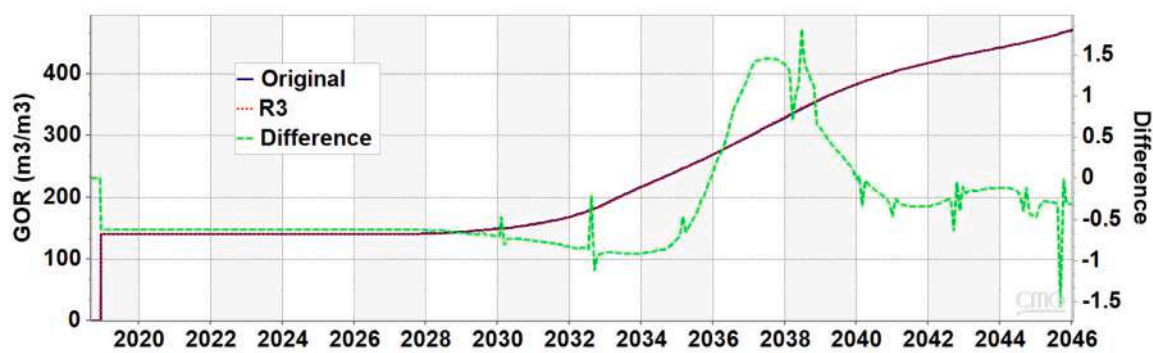
Aiming to evaluate the advantages of our systematic global methodology over the method developed by Alavian et al. (2014), we applied our method to characterize Mixture B (see Table 3). Fig. 5 presents the phase diagrams obtained using the original composition of Mixture B (34 components) and the recharacterizations generated by the methods of Alavian et al. (2014) and our global recharacterization method (see Table 3). The recharacterization mixtures using the method of Lage (2007) presented a better representation of the original fluid than the composition generated assuming the strategy of Alavian et al. (2014), particularly for temperature values larger than 400 °C. Importantly, the global lumping scheme by optimization proposed by Alavian et al. (2014) needed to test 142,506 lumping combinations to determine the best-recharacterized composition, while our method is straightforward. Moreover, the optimization in Alavian et al. (2014) method depends on the choice of weighting factors to weigh the errors in the selected properties, that depends on the application. Therefore, it is not completely systematic.

These results show clear advantages of having an accurate and systematic global recharacterization method that is straightforward in generating a representative reduced fluid composition.

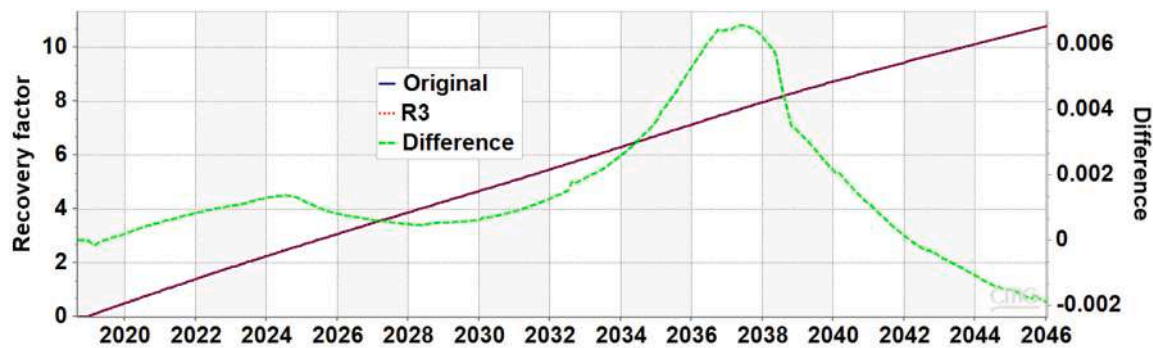
3.2. Error analysis

This section presents the error analysis of PVT results using the recharacterization methods described in Table 3 for Mixture B to quantify the observed advantages of using the systematic global recharacterization method proposed in this work.

Starting with the critical properties' error, Table 4 presents the Euclidean distance calculated for original and recharacterized compositions of Mixture B. The recharacterization using Alavian et al. (2014)

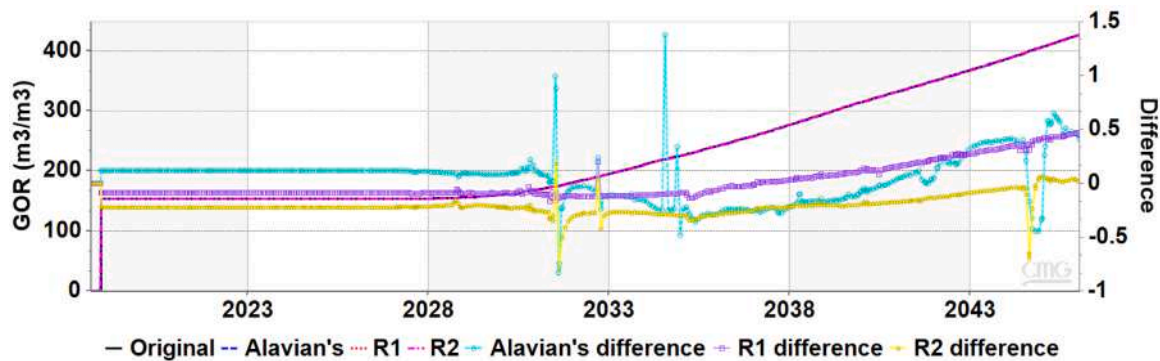


(a)

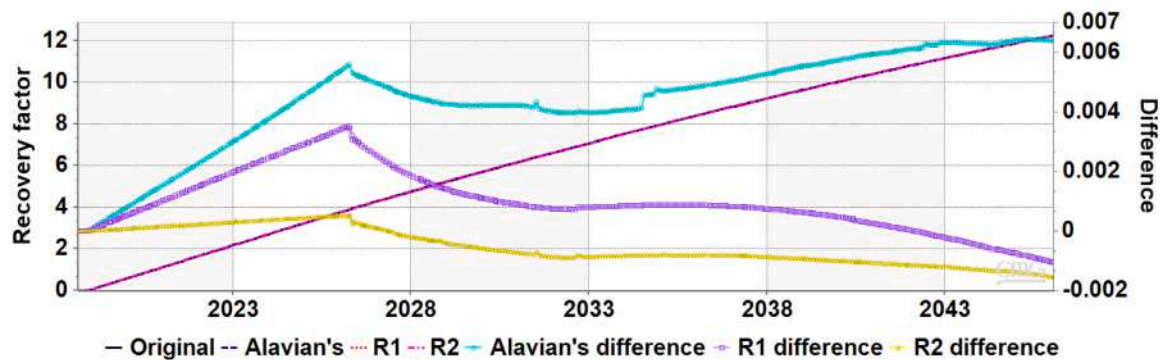


(b)

Fig. 6. Compositional simulation results using Mixture A as reservoir fluid: (a) Gas-oil ratio. (b) Hydrocarbon recovery factor. The secondary axis shows the absolute difference between properties from the recharacterized and original composition.



(a)



(b)

Fig. 7. Compositional simulation results using Mixture B as reservoir fluid: (a) Gas-oil ratio. (b) Hydrocarbon recovery factor. The secondary axis shows the absolute difference between properties from the recharacterized and original composition.

Table 4

Euclidean distance between critical properties points of the original and recharacterized compositions of Mixture B.

Composition	Components	$\frac{T_{c,rec}}{T_{c,o}}$	$\frac{P_{c,rec}}{P_{c,o}}$	$d_{o,rec}$
Original	34	1.0000	1.0000	
Alavian's	9	0.9999	0.9920	0.0080
R1	9	0.9985	1.0001	0.0015
R2	9	0.9993	0.9992	0.0010

Table 5

Vapor fraction, its error, and the molar fraction errors for the Alavian et al.'s recharacterization of Mixture B.

Flash point	β (e_β)	Phase	$e_{c,phase}$	e_c
P1	0.055 (0.008)	liq	0.04	0.082
		vap	0.124	
P2	0.425 (0.003)	liq	0.028	0.038
		vap	0.049	
P3	0.027 (0.003)	liq	0.01	0.019
		vap	0.027	
P4	0.516 (0.009)	liq	0.022	0.023
		vap	0.023	
P5	0.959 (0.012)	liq	0.065	0.067
		vap	0.07	
$\bar{e}_\beta = 0.0068$			$\bar{e}_c = 0.046$	

Table 6

Vapor fraction, its error, and the molar fraction errors for the R1 recharacterization of Mixture B.

Flash point	β (e_β)	Phase	$e_{c,phase}$	e_c
P1	0.059 (0.004)	liq	0.058	0.067
		vap	0.076	
P2	0.427 (0.001)	liq	0.027	0.019
		vap	0.011	
P3	0.023 (0.007)	liq	0.055	0.042
		vap	0.03	
P4	0.527 (0.002)	liq	0.018	0.026
		vap	0.034	
P5	0.947 (0.000)	liq	0.024	0.015
		vap	0.005	
$\bar{e}_\beta = 0.0028$			$\bar{e}_c = 0.034$	

presented the largest value of the Euclidean distance, meaning that this strategy has the largest error regarding the critical points (P_c and T_c) of the original composition, which is quantitative evidence that the method of Alavian et al. (2014) was associated with the worst overall representation of Mixture B, as observed in the phase diagrams.

Tables 5, 6, and 7 present the vapor fraction, its error, and the molar fraction errors for the recharacterizations of Mixture B at the five selected flash conditions (see Fig. 1). R1 and R2 recharacterizations presented smaller \bar{e}_c and \bar{e}_β values than those using the method of Alavian et al. (2014).

3.3. Compositional reservoir simulations

This section presents the investigation of using the original and recharacterized compositions of Mixtures A and B as reservoir fluids in compositional simulations, comparing the predicted reservoir performances.

Fig. 6 presents the gas–oil ratio (Fig. 6(a)) and the hydrocarbon recovery factor (Fig. 6(b)) simulated using the original and R3 recharacterization compositions of Mixture A. Both results behave similarly

Table 7

Vapor fraction, its error, and the molar fraction errors for the R2 recharacterization of Mixture B.

Flash point	β (e_β)	Phase	$e_{c,phase}$	e_c
P1	0.061 (0.002)	liq	0.032	0.032
		vap	0.032	
P2	0.427 (0.001)	liq	0.017	0.012
		vap	0.006	
P3	0.026 (0.005)	liq	0.035	0.023
		vap	0.01	
P4	0.525 (0.000)	liq	0.024	0.021
		vap	0.018	
P5	0.947 (0.000)	liq	0.01	0.007
		vap	0.004	
$\bar{e}_\beta = 0.0016$			$\bar{e}_c = 0.019$	

Table 8

Simulation run-times for recharacterizations of Mixture B.

Composition	Components	Elapsed seconds
Original	34	3539
Alavian's	9	495
R1	9	463
R2	9	474

during the 26 years of gas injection, making it necessary to plot the differences between original and recharacterized mixtures to emphasize the observed behavior. Therefore, the reduction from 41 to 8 components generated the same results but ten times faster, which is a huge computational gain for compositional reservoir simulations.

Fig. 7 presents the compositional simulation results for the gas–oil ratio (Fig. 7(a)) and hydrocarbon recovery (Fig. 7(b)) that used Mixture B compositions as the reservoir fluid. As can be observed, curves are pretty much the same for the original composition with thirty-four components and recharacterized compositions with nine components, presenting also small differences between variables obtained from original and recharacterized mixtures. Note that discrepancies between the two methods are not evident because the reservoir temperature used in the simulations (and typical of such systems) is in the phase diagram zone where the method of Alavian et al. (2014) was able to represent the original behavior of Mixture B fairly (see Fig. 5). However, the method of Alavian et al. (2014) always presented the largest differences between the simulation results using the recharacterized and original compositions, as shown in Fig. 7.

Table 8 gives the average simulation times for using the different compositions of Mixture B, showing that the simulations using recharacterized compositions were 7–7.5 times faster than that with the original composition.

Finally, Figs. 8 and 9 present the global mole fraction of CO₂ in a plane in 2024 and 2046, respectively before and after gas breakthrough. The accurate result for the CO₂ when using simplified fluid models is striking.

Even though Alavian et al. (2014) global lumping method resulted in a recharacterized mixture that gave good reservoir simulation results, it is important to remember that it is more costly and less systematic than the method proposed by this study.

4. Conclusions

This study presented a systematic global recharacterization methodology for reducing the number of components describing a complex mixture. We used two fluid compositions obtained from the literature.

We compared the results for phase diagrams, flashes, and compositional simulations using the original fluid compositions to those

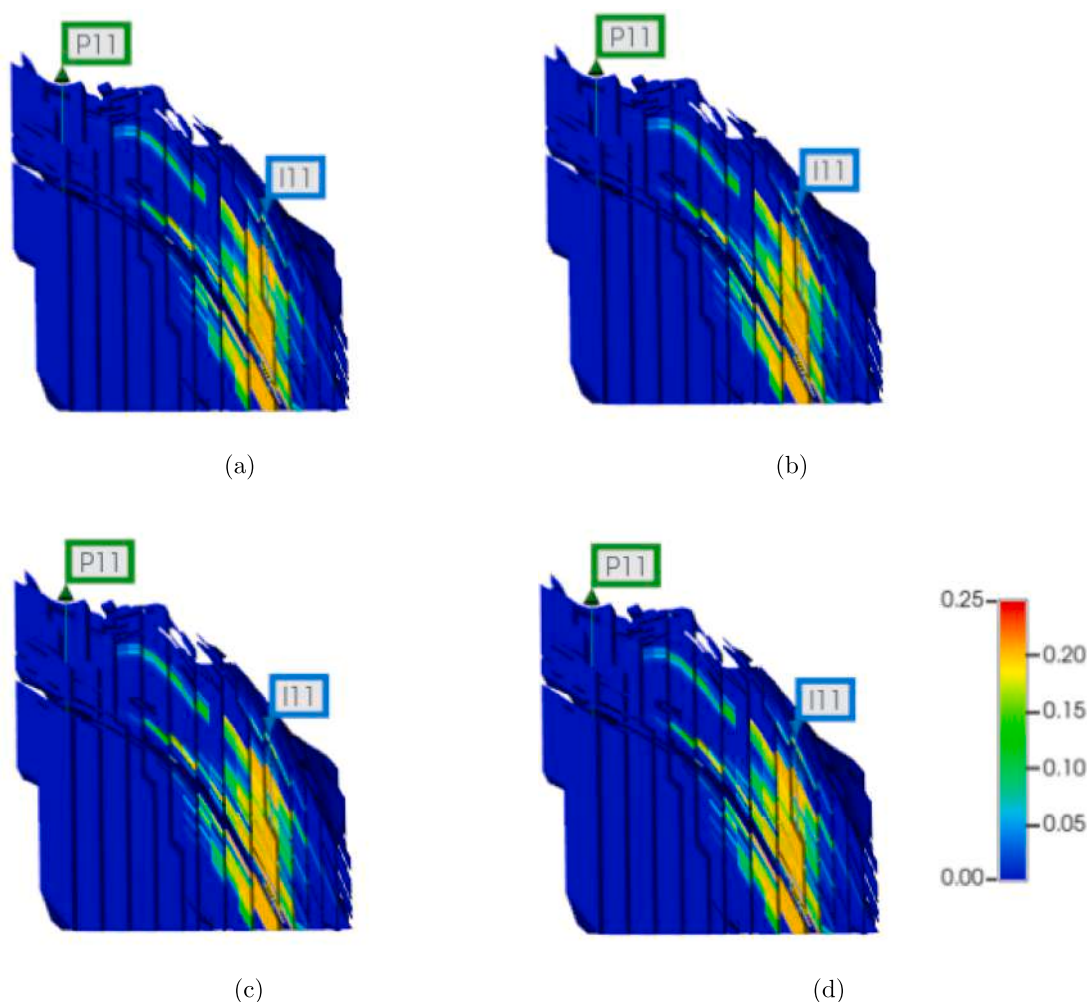


Fig. 8. Mixture B global mole fraction of CO₂ in a reservoir cut at 2024 (before breakthrough): (a) Original composition, (b) Alavian's, (c) R1, and (d) R2.

obtained with their recharacterized compositions. Error analyses of those results showed that our methodology reached results better than those of another one available in the literature.

Importantly, the highly accurate recharacterization method of Lage (2007), which effectively replaces the C_{6+} and C_{7+} fractions by only four or three new pseudo-components, eliminates the need for lumping of important discrete components such as CO₂, methane, and ethane, that significantly improves the simulation results.

Therefore, coupling specific grouping rules for the discrete components and using Lage's method for the C_+ fraction recharacterization provides an accurate fluid description for high-demanding compositional simulations.

CRediT authorship contribution statement

Barbara F. Esteves: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Fellipe C. de Oliveira:** Writing – review & editing, Software, Investigation, Data curation. **Paulo Couto:** Supervision, Resources, Project administration, Funding acquisition. **Paulo L.C. Lage:** Writing – review & editing, Supervision, Software, Resources, Project administration, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Paulo Laranjeira da Cunha Lage reports financial support was provided by Shell Brazil Oil. Paulo Couto reports financial support was provided by Shell Brazil Oil. Barbara de Farias Esteves reports financial support was provided by Shell Brazil Oil. Fellipe Carvalho de Oliveira reports financial support was provided by Shell Brazil Oil. Paulo Laranjeira da Cunha Lage reports financial support was provided by National Council for Scientific and Technological Development. Paulo Couto reports financial support was provided by National Council for Scientific and Technological Development. Paulo Laranjeira da Cunha Lage reports a relationship with Shell Brazil Oil that includes: consulting or advisory. Paulo Couto reports a relationship with Shell Brazil Oil that includes: consulting or advisory and funding grants. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was carried out in association with the ongoing R&D project registered as ANP n° 23209-0, “Caracterização de Fluidos Complexos em Condições de Reservatórios: Avaliação do Comportamento de Fases de Fluidos de Reservatórios Sob a Presença de CO₂” (UFRJ/Shell Brasil/ANP), sponsored by Shell Brasil Petróleo Ltda under the ANP R&D levy as “Compromisso de Investimentos com Pesquisa e Desenvolvimento”. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES)

– Finance Code 001. Paulo L. C Lage also thanks the financial support from CNPq, process number 303845/2022-8. Paulo Couto also acknowledged the financial support through CNPq process number 311143/2022-9.

Appendix A. Correlations for SG and critical properties

Hosseinfar and Jamshidi (2014) proposed a method to predict critical properties, acentric factor, molecular weight, and normal boiling point based on only the molecular mass, assuming an auxiliary relation for the specific gravity. They showed that the proposed method performs more accurately for pure substances and petroleum fractions than conventional methods.

Only the molar masses of the mixtures’ components were available in our study. Thus, we assumed the following auxiliary relation for specific gravity:

$$SG = (a M^b + c^d) \tag{9}$$

where, for petroleum fractions, $a = 0.0089$, $b = 0.70799$, $c = -0.14064$, and $d = 0.13199$. The general form of the correlations for the critical properties and acentric factor is:

$$\Phi(f(SG), M) = (A f(SG)^B M^C + D f(SG)^E M^F) \tag{10}$$

where

$$f(SG) = \sqrt{\frac{3+2 SG}{3-SG}}, \tag{11}$$

and the parameters for Eq. (10) are described for each specific property in Table A.1.

Appendix B. Mixture A

B.1. Original composition

See Table B.1.

B.2. Recharacterized compositions

See Table B.2.

Appendix C. Mixture B

C.1. Original composition

See Table C.1.

C.2. Recharacterized compositions

See Table C.2.

Data availability

The code for calculating Lage’s mixture recharacterization method is available in a public repository.

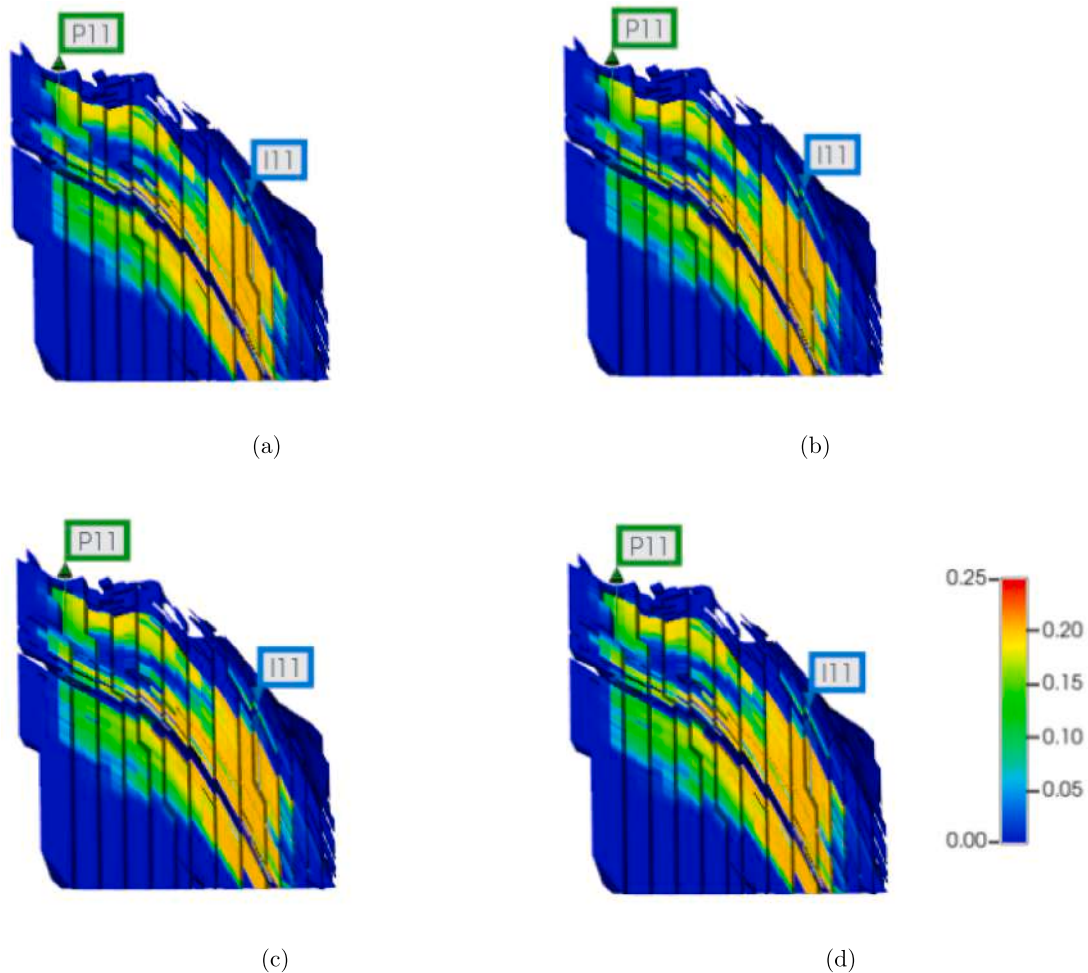


Fig. 9. Mixture B global mole fraction of CO₂ in a reservoir cut at 2046 (after breakthrough): (a) Original composition, (b) Alavian’s, (c) R1, and (d) R2.

Table A.1
Coefficients in Eq. (10) to estimate critical properties for petroleum fractions (Hosseinfar and Jamshidi, 2014).

Φ	A	B	C	D	E	F	G
T_c	-1.60557	-2.97959	-0.02555	6.98661	1.36983	0.04822	2.44263
P_c	0.04555	-0.49263	0.42692	0.58856	7.58044	-0.70954	-5.86206
ω	-0.002972	9.670163	0.20558	0.00126	-1.67098	1.273	0.40017

Table B.1
Original composition of Mixture A. The last column refer to the reference of the critical properties calculation method.

#	Component	M	M_w	P_c	T_c	ω	Reference
1	CO ₂	44.01	10.17	72.80	304.20	0.23	Computer Modelling Group Ltd. (2022)
2	CH ₄	16.04	51.25	45.40	190.60	0.01	Computer Modelling Group Ltd. (2022)
3	C ₂ H ₆	30.07	8.11	48.20	305.40	0.10	Computer Modelling Group Ltd. (2022)
4	C ₃ H ₈	44.10	5.35	41.90	369.80	0.15	Computer Modelling Group Ltd. (2022)
5	<i>i</i> C ₄	58.12	0.01	36.00	408.10	0.18	Computer Modelling Group Ltd. (2022)
6	<i>n</i> C ₄	58.12	0.00	37.50	425.20	0.19	Computer Modelling Group Ltd. (2022)
7	<i>i</i> C ₅	72.15	0.01	33.40	460.40	0.23	Computer Modelling Group Ltd. (2022)
8	<i>n</i> C ₅	72.15	0.01	33.30	469.60	0.25	Computer Modelling Group Ltd. (2022)
9	C ₆	84.00	0.13	32.65	517.20	0.27	Hosseinfar and Jamshidi (2014)
10	m-cyclo-C ₅	84.16	0.08	37.35	532.80	0.23	Computer Modelling Group Ltd. (2022)
11	Benzene	78.11	0.01	48.30	562.10	0.21	Computer Modelling Group Ltd. (2022)
12	Cyclo-C ₆	84.16	0.09	40.20	553.40	0.21	Computer Modelling Group Ltd. (2022)
13	C ₇	96.00	0.44	29.02	549.13	0.30	Hosseinfar and Jamshidi (2014)
14	m-cyclo-C ₆	98.19	0.26	34.26	572.20	0.24	Computer Modelling Group Ltd. (2022)
15	Toluene	92.14	0.02	40.60	591.70	0.26	Computer Modelling Group Ltd. (2022)
16	C ₈	107.00	0.94	26.83	574.18	0.32	Hosseinfar and Jamshidi (2014)
17	C ₂ -benzene	106.17	0.05	35.62	617.15	0.30	Computer Modelling Group Ltd. (2022)
18	m&p-xylene	106.17	0.06	28.24	567.96	0.31	Computer Modelling Group Ltd. (2022)
19	o-xylene	106.16	0.03	36.52	631.15	0.31	Computer Modelling Group Ltd. (2022)
20	C ₉	121.00	1.21	24.79	602.08	0.36	Hosseinfar and Jamshidi (2014)
21	C ₁₀	134.00	1.43	23.31	625.07	0.39	Hosseinfar and Jamshidi (2014)
22	C ₁₁	147.00	1.31	22.05	645.92	0.42	Hosseinfar and Jamshidi (2014)
23	C ₁₂	161.00	1.27	20.85	666.49	0.46	Hosseinfar and Jamshidi (2014)
24	C ₁₃	175.00	1.36	19.78	685.46	0.49	Hosseinfar and Jamshidi (2014)
25	C ₁₄	190.00	1.23	18.74	704.33	0.52	Hosseinfar and Jamshidi (2014)
26	C ₁₅	206.00	1.22	17.72	723.08	0.55	Hosseinfar and Jamshidi (2014)
27	C ₁₆	222.00	0.98	16.78	740.61	0.58	Hosseinfar and Jamshidi (2014)
28	C ₁₇	237.00	0.94	15.96	756.11	0.61	Hosseinfar and Jamshidi (2014)
29	C ₁₈	251.00	0.95	15.25	769.85	0.64	Hosseinfar and Jamshidi (2014)
30	C ₁₉	263.00	0.84	14.67	781.14	0.66	Hosseinfar and Jamshidi (2014)
31	C ₂₀	275.00	0.72	14.13	792.02	0.68	Hosseinfar and Jamshidi (2014)
32	C ₂₁	291.00	0.67	13.44	805.94	0.71	Hosseinfar and Jamshidi (2014)
33	C ₂₂	305.00	0.62	12.87	817.63	0.73	Hosseinfar and Jamshidi (2014)
34	C ₂₃	318.00	0.56	12.38	828.11	0.75	Hosseinfar and Jamshidi (2014)
35	C ₂₄	331.00	0.55	11.91	838.25	0.77	Hosseinfar and Jamshidi (2014)
36	C ₂₅	346.00	0.52	11.40	849.58	0.80	Hosseinfar and Jamshidi (2014)
37	C ₂₆	359.00	0.47	10.98	859.09	0.82	Hosseinfar and Jamshidi (2014)
38	C ₂₇	374.00	0.46	10.52	869.74	0.84	Hosseinfar and Jamshidi (2014)
39	C ₂₈	388.00	0.45	10.12	879.38	0.86	Hosseinfar and Jamshidi (2014)
40	C ₂₉	402.00	0.44	9.74	888.76	0.88	Hosseinfar and Jamshidi (2014)
41	C ₃₀₊	652.69	4.79	5.33	1025.59	1.18	Hosseinfar and Jamshidi (2014)

Table B.2
Recharacterized compositions of Mixture A. The last column refer to the reference of the critical properties calculation method.

#	Component	M	M_w	P_c	T_c	ω	Reference
R1							
1	CO ₂	44.01	10.17	72.80	304.20	0.23	Computer Modelling Group Ltd. (2022)
2	CH ₄	16.04	51.25	45.40	190.60	0.01	Computer Modelling Group Ltd. (2022)
3	C ₂ - C ₃	35.65	13.47	45.70	343.69	0.12	Kay (1936)
4	C ₄ - C ₅	65.64	0.03	34.74	443.43	0.21	Kay (1936)
5	Pseudo 1	137.28	10.61	22.97	630.51	0.40	Hosseinfar and Jamshidi (2014)
6	Pseudo 2	297.54	9.52	13.17	811.46	0.72	Hosseinfar and Jamshidi (2014)
7	Pseudo 3	650.16	4.95	5.35	1024.42	1.18	Hosseinfar and Jamshidi (2014)
R2							
1	CO ₂	44.01	10.17	72.80	304.20	0.23	Computer Modelling Group Ltd. (2022)
2	CH ₄	16.04	51.25	45.40	190.60	0.01	Computer Modelling Group Ltd. (2022)
3	C ₂ - C ₃	35.65	13.47	45.70	343.69	0.12	Kay (1936)
4	<i>i</i> C ₄	58.12	0.01	36.00	408.10	0.18	Computer Modelling Group Ltd. (2022)
5	<i>n</i> C ₄	58.12	0.00	37.50	425.20	0.19	Computer Modelling Group Ltd. (2022)
6	<i>i</i> C ₅	72.15	0.01	33.40	460.40	0.23	Computer Modelling Group Ltd. (2022)

(continued on next page)

Table B.2 (continued).

#	Component	M	M_w	P_c	T_c	ω	Reference
7	nC_5	72.15	0.01	33.30	469.60	0.25	Computer Modelling Group Ltd. (2022)
8	Pseudo 1	137.28	10.61	22.97	630.51	0.40	Hosseinfar and Jamshidi (2014)
9	Pseudo 2	297.54	9.52	13.17	811.46	0.72	Hosseinfar and Jamshidi (2014)
10	Pseudo 3	650.16	4.95	5.35	1024.42	1.18	Hosseinfar and Jamshidi (2014)
R3							
1	CO ₂	44.01	10.17	72.80	304.20	0.23	Computer Modelling Group Ltd. (2022)
2	CH ₄	16.04	51.25	45.40	190.60	0.01	Computer Modelling Group Ltd. (2022)
3	C ₂ H ₆	30.07	8.11	48.20	305.40	0.10	Computer Modelling Group Ltd. (2022)
4	C ₃ H ₈	44.10	5.35	41.90	369.80	0.15	Computer Modelling Group Ltd. (2022)
5	C ₄ – C ₅	65.64	0.03	34.74	443.43	0.21	Kay (1936)
6	Pseudo 1	137.28	10.61	22.97	630.51	0.40	Hosseinfar and Jamshidi (2014)
7	Pseudo 2	297.54	9.52	13.17	811.46	0.72	Hosseinfar and Jamshidi (2014)
8	Pseudo 3	650.16	4.95	5.35	1024.42	1.18	Hosseinfar and Jamshidi (2014)
R4							
1	CO ₂	44.01	10.17	72.80	304.20	0.23	Computer Modelling Group Ltd. (2022)
2	CH ₄	16.04	51.25	45.40	190.60	0.01	Computer Modelling Group Ltd. (2022)
3	C ₂ H ₆	30.07	8.11	48.20	305.40	0.10	Computer Modelling Group Ltd. (2022)
4	C ₃ H ₈	44.10	5.35	41.90	369.80	0.15	Computer Modelling Group Ltd. (2022)
5	C ₄ – C ₅	65.64	0.03	34.74	443.43	0.21	Kay (1936)
6	Pseudo 1	116.94	6.33	25.33	594.38	0.35	Hosseinfar and Jamshidi (2014)
7	Pseudo 2	216.06	9.41	17.12	734.24	0.57	Hosseinfar and Jamshidi (2014)
8	Pseudo 3	352.72	4.54	11.18	854.53	0.81	Hosseinfar and Jamshidi (2014)
8	Pseudo 4	652.67	4.80	5.33	1025.58	1.18	Hosseinfar and Jamshidi (2014)
R5							
1	CO ₂	44.01	10.17	72.80	304.20	0.23	Computer Modelling Group Ltd. (2022)
2	CH ₄	16.04	51.25	45.40	190.60	0.01	Computer Modelling Group Ltd. (2022)
3	C ₂ H ₆	30.07	8.11	48.20	305.40	0.10	Computer Modelling Group Ltd. (2022)
4	C ₃ H ₈	44.10	5.35	41.90	369.80	0.15	Computer Modelling Group Ltd. (2022)
5	C ₄ – C ₆	28.45	0.33	35.68	524.16	0.24	Kay (1936)
6	Pseudo 1	140.81	10.55	22.62	636.23	0.41	Hosseinfar and Jamshidi (2014)
7	Pseudo 2	299.70	9.29	13.08	813.26	0.72	Hosseinfar and Jamshidi (2014)
8	Pseudo 3	650.31	4.95	5.35	1024.49	1.18	Hosseinfar and Jamshidi (2014)
R6							
1	CO ₂	44.01	10.17	72.80	304.20	0.23	Computer Modelling Group Ltd. (2022)
2	CH ₄	16.04	51.25	45.40	190.60	0.01	Computer Modelling Group Ltd. (2022)
3	C ₂ H ₆	30.07	8.11	48.20	305.40	0.10	Computer Modelling Group Ltd. (2022)
4	C ₃ H ₈	44.10	5.35	41.90	369.80	0.15	Computer Modelling Group Ltd. (2022)
5	C ₄ – C ₆	28.45	0.33	35.68	524.16	0.24	Kay (1936)
6	Pseudo 1	121.63	6.45	24.71	603.25	0.36	Hosseinfar and Jamshidi (2014)
7	Pseudo 2	219.89	9.12	166.90	738.36	0.58	Hosseinfar and Jamshidi (2014)
8	Pseudo 3	354.19	4.41	11.13	855.60	0.81	Hosseinfar and Jamshidi (2014)
9	Pseudo 4	652.69	4.80	5.33	1025.58	1.18	Hosseinfar and Jamshidi (2014)

Table C.1

Original composition of Mixture B.

#	Components	% (mol)	M (g/mol)	P_c (atm)	T_c (K)	ω
1	N ₂	0.072	28.01	33.54	126.20	0.04
2	C ₁	38.249	16.04	45.39	190.56	0.01
3	CO ₂	0.160	44.01	72.78	304.12	0.23
4	C ₂	12.461	30.07	48.08	305.32	0.10
5	C ₃	8.729	44.10	41.92	369.83	0.15
6	<i>i</i> C ₄	1.036	58.12	35.92	407.85	0.19
7	<i>n</i> C ₄	4.198	58.12	37.46	425.12	0.20
8	<i>i</i> C ₅	1.148	72.15	33.37	460.39	0.23
9	<i>n</i> C ₅	2.102	72.15	33.26	469.70	0.25
10	C ₆	3.023	82.42	33.34	513.35	0.24
11	C ₇	2.920	96.10	30.90	550.43	0.28
12	C ₈	2.826	108.94	28.66	579.80	0.31
13	C ₉	2.462	122.09	26.43	607.62	0.35
14	C ₁₀	2.189	135.01	24.51	632.28	0.39
15	C ₁₁	1.941	147.85	22.83	654.64	0.43
16	C ₁₂	1.717	160.59	21.36	675.04	0.47
17	C ₁₃	1.517	173.24	20.07	693.78	0.51
18	C ₁₄	1.339	185.78	18.92	711.06	0.54
19	C ₁₅	1.181	198.22	17.91	727.08	0.58
20	C ₁₆	1.042	210.55	17.00	742.00	0.61
21	C ₁₇	0.919	222.77	16.19	755.92	0.65

(continued on next page)

Table C.1 (continued).

#	Components	% (mol)	M (g/mol)	P_c (atm)	T_c (K)	ω
22	C ₁₈	0.811	234.88	15.46	768.98	0.68
23	C ₁₉	0.716	246.87	14.80	781.24	0.71
24	C ₂₀	0.632	258.75	14.21	792.81	0.75
25	C ₂₁	0.559	270.52	13.67	803.73	0.78
26	C ₂₂	0.494	282.18	13.17	814.08	0.81
27	C ₂₃	0.438	293.73	12.72	823.91	0.84
28	C ₂₄	0.388	305.17	12.31	833.26	0.87
29	C ₂₅	0.345	316.50	11.93	842.16	0.90
30	C ₂₆	0.306	327.73	11.58	850.67	0.93
31	C ₂₇	0.273	338.86	11.26	858.80	0.95
32	C ₂₈	0.243	349.89	10.96	866.59	0.98
33	C ₂₉	0.217	360.82	10.68	874.07	1.01
34	C ₃₀₊	3.348	496.34	8.30	951.02	1.30

Table C.2

Recharacterized compositions of Mixture B.

#	Components	% (mol)	M (g/mol)	P_c (atm)	T_c (K)	ω
Alavian's						
1	N ₂	0.072	28.01	33.54	126.20	0.04
2	C ₁	38.249	16.04	45.39	190.56	0.01
3	CO ₂	0.160	44.01	72.78	304.12	0.23
4	C ₂ – C ₃	21.190	35.89	45.19	333.27	0.12
5	C ₄ – C ₆	11.507	68.59	34.89	459.98	0.22
6	C ₇ – C ₁₁	12.338	119.55	24.98	599.36	0.35
7	C ₁₂ – C ₁₅	5.754	177.67	19.59	688.92	0.49
8	C ₁₆ – C ₂₃	5.610	245.49	15.52	764.52	0.63
9	C ₂₄ – C ₃₀₊	5.120	444.01	8.70	915.48	0.94
R1						
1	CO ₂	0.160	44.01	72.78	304.12	0.23
2	N ₂ – C ₁	38.320	16.06	45.37	190.44	0.01
3	C ₂	12.461	30.07	48.08	305.32	0.10
4	C ₃	8.729	44.10	41.92	369.83	0.15
5	C ₄ – C ₆	11.507	68.47	35.05	458.19	0.23
6	Pseudo 1	10.347	109.51	26.42	579.48	0.33
7	Pseudo 2	10.911	186.25	18.99	699.74	0.51
8	Pseudo 3	4.176	303.32	12.94	816.25	0.73
9	Pseudo 4	3.389	495.89	7.63	945.97	1.00
R2						
1	CO ₂	0.160	44.01	72.78	304.12	0.23
2	N ₂ – C ₁	38.320	16.06	45.37	190.44	0.01
3	C ₂	12.461	30.07	48.08	305.32	0.10
4	C ₃	8.729	44.10	41.92	369.83	0.15
5	C ₄ – C ₅	8.484	63.49	35.68	438.83	0.22
6	Pseudo 1	11.431	97.23	28.73	552.11	0.30
7	Pseudo 2	12.397	176.35	19.69	687.22	0.49
8	Pseudo 3	4.619	298.56	13.13	812.31	0.72
9	Pseudo 4	3.399	495.77	7.63	945.90	1.00

References

- Alavian, Sayyed Ahmad, Whitson, Curtis Hays, Martinsen, Sissel O., 2014. Global component lumping for EOS calculations. In: SPE Annual Technical Conference and Exhibition, All Days: SPE-170912-MS. <http://dx.doi.org/10.2118/170912-MS>.
- Botechia, V.E., Correia, M.G., Rios, V.S., Davolio, A., Santos, S.M.G., Hohen-dorff Filho, J.C.V., Schiozer, D.J., 2023. Unisim-iv: Benchmark proposal for light oil carbonate reservoir with high co2 content. *Braz. J. Pet. Gas* 16 (4).
- Chicralla, Felipe C., Lage, Paulo L.C., Secchi, Argimiro R., 2019. Quadrature algorithms for phase equilibrium of continuous mixtures. *Braz. J. Chem. Eng.* 36 (3), 1303–1318.
- Christie, M.A., Blunt, M.J., 2001. Tenth spe comparative solution project: A comparison of upscaling techniques. *SPE Reserv. Eval. Eng.* (ISSN: 10946470) 4 (4), 308–316 – 316.
- Coats, K.H., Smart, G.T., 1986. Application of a Regression-Based EOS PVT Program to Laboratory Data. *SPE Reserv. Eng.* (ISSN: 0885-9248) 1 (03), 277–299. <http://dx.doi.org/10.2118/11197-PA>.
- Computer Modelling Group Ltd., 2022. Cmg. <https://hadoop.apache.org>.
- Costa, Daniela, Coelho, Anete, Santos, Davi, Magalhães, Shayane, 2022. RELATÓRIO PVT Caracterização de Fluidos. Technical Report FLUID-02-001, Laboratório de Recuperação Avançada de Petróleo (LRAP) UFRJ - Universidade Federal do Rio de Janeiro.
- Danesh, Ali, Xu, Dong-hai, Todd, Adrian C., 1992. A grouping method to optimize oil description for compositional simulation of gas-injection processes. *SPE Reserv. Eng.* 7 (03), 343–348. <http://dx.doi.org/10.2118/20745-PA>.
- Gross, Joachim, Sadowski, Gabriele, 2001. Perturbed-chain saft: An equation of state based on a perturbation theory for chain molecules. *Ind. Eng. Chem. Res.* 40 (4), 1244–1260. <http://dx.doi.org/10.1021/ie0003887>.
- He, Jincong, Durlafsky, Louis J., 2014. Reduced-order modeling for compositional simulation by use of trajectory piecewise linearization. *SPE J.* (ISSN: 1086-055X) 19 (05), 858–872. <http://dx.doi.org/10.2118/163634-PA>.
- Honami, Y., Arihara, N., Yazawa, N., 2000. Accuracy and efficiency evaluation of EOS computation methods for compositional simulation. In: SPE Asia Pacific Conference on Integrated Modelling for Asset Management, All Days: SPE-59428-MS. <http://dx.doi.org/10.2118/59428-MS>.
- Hosseinfar, Pouya, Jamshidi, Saeid, 2014. Development of a new generalized correlation to characterize physical properties of pure components and petroleum fractions. *Fluid Phase Equilib.* (ISSN: 0378-3812) 363, 189–198. <http://dx.doi.org/10.1016/j.fluid.2013.11.043>.
- Joergensen, M., Stenby, E.H., 1995. Optimization of pseudo-component selection for compositional studies of reservoir fluids. In: SPE Annual Technical Conference and Exhibition, SPE-30789-MS. <http://dx.doi.org/10.2118/30789-MS>.
- John, Volker, Thein, Ferdinand, 2012. On the efficiency and robustness of the core routine of the quadrature method of moments (qmom). *Chem. Eng. Sci.* 75, 327–333.
- Kay, W., 1936. Gases and vapors at high temperature and pressure - density of hydrocarbon. *Ind. Eng. Chem.* 28 (9), 1014–1019. <http://dx.doi.org/10.1021/ie50321a008>.
- Lage, Paulo L.C., 2007. The quadrature method of moments for continuous thermodynamics. *Comput. Chem. Eng.* (ISSN: 0098-1354) 31 (7), 782–799. <http://>

- dx.doi.org/10.1016/j.compchemeng.2006.08.005, <https://www.sciencedirect.com/science/article/pii/S0098135406002146>.
- Lage, Paulo L.C., 2024. Lage's method for mixture re-characterization. Gitlab public project, https://gitlab.com/paulolage/lage_method.
- Lee, Byung Ik, Kesler, Michael G., 1975. A generalized thermodynamic correlation based on three-parameter corresponding states. *AIChE J.* 21 (3), 510–527. <http://dx.doi.org/10.1002/aic.690210313>.
- Lohrenz, John, Bray, Bruce G., Clark, Charles R., 1964. Calculating viscosities of reservoir fluids from their compositions. *J. Pet. Technol.* 16 (10), 1171–1176.
- Lolley, C.S., Richardson, W.C., 1998. Minimum pseudocomponent requirements for compositional thermal simulation of. In: *SPE Improved Oil Recovery Conference*, All Days: SPE-39640-MS. <http://dx.doi.org/10.2118/39640-MS>.
- Mello, S.F., Ligerio, E.L., Scanavini, H.F., Schiozer, D.J., 2011. Influence of lumping and equation of state tuning methods on the sub-salt reservoirs simulation. In: *SPE Europec featured at EAGE Conference and Exhibition*, SPE-143041-MS. <http://dx.doi.org/10.2118/143041-MS>.
- Moghadamzadeh, Hamidreza, Maghsoodloordad, Hojjatolla, Zarabpour, Adel, Hemmati, Ali, Shahsavari, Shahin, Shahsavari, Shadab, 2013. Impact of lumping techniques for fluid characterization in gas condensate reservoir. *Aust. J. Basic Appl. Sci.* (ISSN: 1991-8178) 7 (1), 320–333.
- Montel, F., Gouel, P.L., 1984. A new lumping scheme of analytical data for compositional studies. In: *SPE Annual Technical Conference and Exhibition*, All Days: SPE-13119-MS. <http://dx.doi.org/10.2118/13119-MS>.
- Newley, T.M.J., Merrill, Jr., R.C., 1991. Pseudocomponent selection for compositional simulation. *SPE Reserv. Eng.* (ISSN: 0885-9248) 6 (04), 490–496. <http://dx.doi.org/10.2118/19638-PA>.
- Oliveira, Felipe C., Lage, Paulo L.C., Couto, Paulo, 2023. Comparison of lumping methods to adaptive characterization in representing phase equilibria of hydrocarbon mixtures using pc-saft eos. *Chem. Eng. Sci.* (ISSN: 0009-2509) 119066. <http://dx.doi.org/10.1016/j.ces.2023.119066>.
- Pedersen, Karen Schou, Fredenslund, Aage, 1987. An improved corresponding states model for the prediction of oil and gas viscosities and thermal conductivities. *Chem. Eng. Sci.* 42 (1), 182–186.
- Pedersen, Karen Schou, Fredenslund, Aage, Thomassen, Per, 1989. *Properties of Oils and Natural Gases*, vol. 5, Butterworth-Heinemann.
- Pedersen, Karen Schou, Thomassen, Per, Fredenslund, Aage, 1985. Thermodynamics of petroleum mixtures containing heavy hydrocarbons. 3. efficient flash calculation procedures using the srk equation of state. *Ind. Eng. Chem. Process. Des. Dev.* 24 (4), 948–954. <http://dx.doi.org/10.1021/i200031a009>.
- Pedersen, Karen Schou, Thomassen, Per, Fredenslund, Aage, 1988. On the dangers of tuning equation of state parameters. *Chem. Eng. Sci.* (ISSN: 0009-2509) 43 (2), 269–278. [http://dx.doi.org/10.1016/0009-2509\(88\)85039-5](http://dx.doi.org/10.1016/0009-2509(88)85039-5).
- Peng, Ding-Yu, Robinson, Donald B., 1976. A new two-constant equation of state. *Ind. Eng. Chem. Fundam.* 15 (1), 59–64.
- Rastegar, Reza, Jessen, Kristian, 2009. Lumping and delumping for integrated compositional modeling. In: *SPE Annual Technical Conference and Exhibition*, All Days: SPE-125017-MS. <http://dx.doi.org/10.2118/125017-MS>.
- Twu, Chong H., 1984. An internally consistent correlation for predicting the critical properties and molecular weights of petroleum and coal-tar liquids. *Fluid Phase Equilib.* (ISSN: 0378-3812) 16 (2), 137–150. [http://dx.doi.org/10.1016/0378-3812\(84\)85027-X](http://dx.doi.org/10.1016/0378-3812(84)85027-X).
- Wang, Jun, Jia, Chun-Sheng, Li, Chang-Jun, Peng, Xiao-Long, Zhang, Lie-Hui, Liu, Jian-Yi, 2019. Thermodynamic properties for carbon dioxide. *ACS Omega* 4 (21), 19193–19198.
- Whitson, Curtis H., 1983. Characterizing hydrocarbon plus fractions. *Soc. Pet. Eng. J.* (ISSN: 0197-7520) 23 (04), 683–694. <http://dx.doi.org/10.2118/12233-PA>.
- Zick Technologies, 2024. Phazecom. <https://www.zicktech.com/phazecom.html>.