



Towards a predictive Cubic Plus Association equation of state

Velho, Pedro; Liang, Xiaodong; Macedo, Eugénia A.; Gómez, Elena; Kontogeorgis, Georgios M.

Published in:
Fluid Phase Equilibria

Link to article, DOI:
[10.1016/j.fluid.2021.113045](https://doi.org/10.1016/j.fluid.2021.113045)

Publication date:
2021

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Velho, P., Liang, X., Macedo, E. A., Gómez, E., & Kontogeorgis, G. M. (2021). Towards a predictive Cubic Plus Association equation of state. *Fluid Phase Equilibria*, 540, Article 113045.
<https://doi.org/10.1016/j.fluid.2021.113045>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

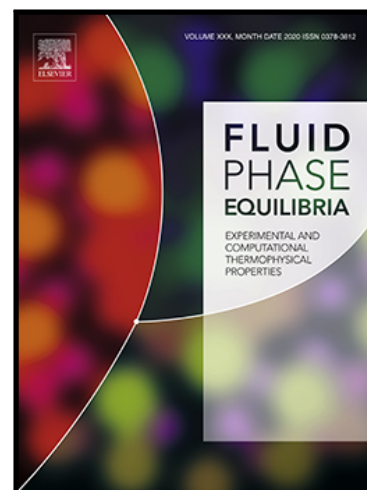
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Journal Pre-proof

Towards a predictive Cubic Plus Association equation of state

Pedro Velho, Xiaodong Liang, Eugénia A. Macedo, Elena Gómez,
Georgios M. Kontogeorgis

PII: S0378-3812(21)00107-2
DOI: <https://doi.org/10.1016/j.fluid.2021.113045>
Reference: FLUID 113045



To appear in: *Fluid Phase Equilibria*

Received date: 29 December 2020
Revised date: 14 March 2021
Accepted date: 2 April 2021

Please cite this article as: Pedro Velho, Xiaodong Liang, Eugénia A. Macedo, Elena Gómez, Georgios M. Kontogeorgis, Towards a predictive Cubic Plus Association equation of state, *Fluid Phase Equilibria* (2021), doi: <https://doi.org/10.1016/j.fluid.2021.113045>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 Published by Elsevier B.V.

Towards a predictive Cubic Plus Association equation of state

Pedro Velho^{a,b}, Xiaodong Liang^a, Eugnia A. Macedo^b, Elena Gmez^b, Georgios M. Kontogeorgis^{a,*}

^aCenter for Energy Resources Engineering (CERE), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, Søltofts Plads, DK2800, Kongens Lyngby, Denmark.

^bAssociate Laboratory of Separation and Reaction Engineering Laboratory of Catalysis and Materials (LSRE-LCM), Department of Chemical Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, s/n, 4200-465, Porto, Portugal.

Abstract. The Cubic Plus Association (CPA) equation of state combines the Soave-Redlich-Kwong equation of state with the association term from the Wertheim theory. To describe the phase behaviour of chemical compounds, three parameters are required for non-associating compounds and five are needed for associating compounds. However, these parameters are hard to find for some chemical compounds, for example, ionic liquids, pharmaceuticals, polymers and other heavy compounds. Being so, there is a need to predict them. Using a database with 197 chemical compounds of different families, correlations for parameters and properties were investigated between groups of similar species. Then, four different methods, which used correlated and fitted parameters, were used in predicting vapour pressure and liquid density from $0.5 \cdot T_c$ to $0.95 \cdot T_c$ and the results were compared with data from the DIPPR database. Moreover, in a database with 219 binary interaction parameters (k_{ij}), the exponents n of the Mie potential function were calculated using the Hudson-McCoubrey theory and following two different methods. Overall, by exploring the common behaviour of the properties and parameters of similar compounds, the predictive ability of the CPA EoS model was successfully enlarged.

Keywords: CPA EoS, Predictive model, Hudson-McCoubrey, Mie potential function.

* Corresponding author. Phone number: +45 45 25 28 59. E-mail adress: gk@kt.dtu.dk

1 Introduction

The Cubic Plus Association (CPA) equation of state (EoS) [1] combines the Soave-Redlich-Kwong (SRK) [2] equation of state with the association term from the Wertheim theory. This model is considered an improvement of the cubic equations of state because of its ability to explicitly take hydrogen bonding [3], i.e., association, into account. To do so, specially for systems with associating chemical compounds and complex phase equilibria, the CPA EoS requires five parameters, which are normally found in the literature. However, the CPA EoS parameters of ionic liquids, pharmaceuticals, polymers and other heavy compounds are generally unknown, which creates the need of fitting the parameters to experimental data, such as saturation pressure and liquid density [4]. When these data are unavailable, it may also be required to perform long and expensive laboratory procedures. Being so, one easier way of estimating the CPA EoS parameters would be predicting them

based on known chemical and physical properties. Therefore, the motivation of this work was to develop a more predictive version of the CPA EoS.

For non-associating compounds, i.e., those which do not have hydrogen bonds, the CPA EoS is only described by the SRK EoS, which requires three parameters for pure compounds (a_0 , b and c_0). These parameters have been shown to follow well-defined behaviours for some chemical families, when plotted against the van der Waals volume (V_w) [5], i.e., the volume occupied by each mole of the compound, which is considered impenetrable to other molecules at ordinary temperatures [6]. However, for associating compounds, the Wertheim's theory adds two more parameters (β and ϵ), for which no clear trends have been found so far.

Some previous studies have already focused on predicting the CPA EoS parameters. Palma *et al.* [7], for example, modified the CPA EoS using an extended version of the Mathias-Copeman function for the energy parameter (a) [8]. In their studies, well-defined tendencies were reported for the energy and co-volume parameters with the van der Waals volume for alcohols and diols. Moreover, good correlations were also found for the binary interaction parameter (k_{ij}) with the molecular mass for monoethylene glycol and light alkanes.

Ali [9], on the other hand, found well-defined trends for the energy term parameter (c_1) with the molecular mass of hydrocarbons, for the co-volume (b) with the molecular mass of hydrocarbons and noted a linear correlation between the co-volume and the energy term parameter (a_0).

In another approach to a more predictive model, Hansen [5] supposed second degree trends between the CPA parameters and the van der Waals volume and correlated the trends directly to saturation pressure and liquid density data. The new method, named "trend fit method", was applied to alcohols, primary amines and carboxylic acids and successfully predicted the CPA parameters in all the studied families [5].

Another approach towards a more predictive model was developed by Schlaikjer [10], who obtained good modelling results by predicting the association parameters and keeping them constant for elements of the same chemical family [5]. To what concerns trends, Schlaikjer [10] found a linear first-degree correlation between the co-volume (b) and the van der Waals volume (V_w). For the energy term parameter a_0 , a good second degree relationship was found against the van der Waals volume. In Schlaikjer's [10] analysis, a set of compounds including alkanols, alkanes, esters, ethers, glycols, ketones and phenols was considered and all the fluorine-containing compounds had to be disregarded in the correlations, so more extensive studies were advised to understand their atypical behaviour.

Coutinho *et al.* [11] used the Hudson-McCoubrey method to predict the binary interaction parameters (k_{ij}) of the cubic EoS and fitted the exponent of the Mie potential function

(n). Moreover, some trends of the exponent n with measures of system size asymmetry were found.

However, all these studies were based on small databases and only very preliminary results and conclusions were obtained. Therefore, the objective of this work was to develop a more predictive version of the CPA EoS, with the particular aims of reducing the number of required parameters, of decreasing the need of experimental procedures and of allowing for a better characterisation of a larger number of systems.

2 Model presentation

2.1 The Cubic Plus Association equation of state

The Cubic Plus Association equation of state (CPA EoS) is presented in equation 1 [3].

$$P = \frac{R_g T}{V_m - b} - \frac{a(T)}{V_m(V_m + b)} - \frac{1}{2} \frac{R_g T}{V_m} \left[1 + \frac{1}{V_m} \frac{\partial \ln g(\rho)}{\partial \rho} \right] \sum_i x_i \sum_{A_i} (1 - X_{A_i}) \quad (1)$$

where P is the pressure, T is the temperature, R_g is the ideal gas constant, a is an energy parameter, V_m is the molecular volume, b is the co-volume, g is the radial distribution function, ρ is the molar density, x_i is the mole fraction of the component i and X_{A_i} is related to the association strength. The summations should be performed over all bonding sites and for all components [5].

The version of the CPA EoS presented by equation 1 is often known as simplified Cubic Plus Association EoS, since it is the result of two important simplifications [10]: having used the association term in the form proposed by Michelsen *et al.* [12] and having used the radial function suggested by Kontogeorgis *et al.* [13] and Elliott *et al.* [14].

The energy parameter (a) accounts for the van der Waals interactions between molecules, following equation 2 [15].

$$a(T) = a_0 \left[1 + c_1 (1 - \sqrt{T_r}) \right]^2 \quad (2)$$

where a_0 and c_1 are two pure component parameters, and T_r is the reduced temperature.

The co-volume parameter (b) is a measure of the molecular volume and can be calculated if the free volume is subtracted from the total volume. The association parameters (ϵ and β), only appearing in the association term, are related to the bonding strength of the two association sites: ϵ refers to the energy of association itself and the product $b \cdot \beta$ is the volume of association [5].

The terminology of bonding sites and association schemes, i.e., the number and type of association sites [3], follows Huang and Radosz [16], so X_{A_i} can be calculated using equation:

$$X_{A_i} = \frac{1}{1 + \rho \sum_j x_j \sum_{B_j} X_{B_j} \Delta_{A_i B_j}} \quad (3)$$

where B_j stands for an association site B in molecule j and $\Delta_{A_i B_j}$ for the association strength, calculated for the CPA EoS by equation:

$$\Delta_{A_i B_j} = g(\rho) \left[\exp\left(\frac{\epsilon^{A_i B_j}}{R_g T}\right) - 1 \right] b_{ij} \beta^{A_i B_j} \quad (4)$$

where b_{ij} is the cross co-volume, $\epsilon^{A_i B_j}$ is the cross-association energy parameter and $\beta^{A_i B_j}$ is the cross-association volume parameter [15]. The superscripts of the last two parameters are normally omitted for pure compounds, so the parameters will be referred to as ϵ and β , respectively.

2.2 Mixing rules for the physical term

The establishment of mixing rules is required in order to extend the CPA EoS to mixtures and, typically, the van der Waals one fluid mixing rules (vdW1f) are used, which calculate the energy parameter using equation 5 [3].

$$a = \sum_i \sum_j x_i x_j a_{ij} \quad (5)$$

where a_{ij} is the cross energy parameter, which normally is obtained by the geometric mean of the energy parameters, given by:

$$a_{ij} = \sqrt{a_i a_j} (1 - k_{ij}) \quad (6)$$

where k_{ij} is the binary interaction parameter, an adjustable constant normally obtained from experimental phase equilibria data [5].

On the other hand, the vdW1f mixing rules state that the co-volume can be calculated by:

$$b = \sum_i \sum_j x_i x_j b_{ij} \quad (7)$$

where the cross co-volume (b_{ij}) is commonly obtained via an arithmetic mean, equation 8 [11].

$$b_{ij} = \frac{b_i + b_j}{2} \quad (8)$$

2.3 Combining rules for the physical term

The interaction parameters in the cubic equations of state are significantly affected by the chosen combining rules and this is particularly important for the cross-energy parameter [3], so combining rules have been extensively studied [17, 18].

Reed [19] showed that combining rules could be improved via the intermolecular attraction potential by accounting for differences in the chemical compounds' molar volumes and ionisation potentials, i.e., the energy required to remove the most exterior electron.

Hudson and McCoubrey [20] reformulated Reed's [19] equations using the Lennard-Jones (LJ) 12:6 potential and based on the London theory of dispersion. To calculate the unlike energy parameter between two central molecules (ϵ_{ij}), Hudson and McCoubrey [20] wrote:

$$\epsilon_{ij} = \frac{2\sqrt{I_i I_j}}{I_i + I_j} \left[\frac{\sqrt{\sigma_i \sigma_j}}{\sigma_{ij}} \right]^6 \sqrt{\epsilon_i^{\text{LJ}} \epsilon_j^{\text{LJ}}} \quad (9)$$

where I is the ionisation potential, ϵ^{LJ} is the depth of the potential well or energy parameter, σ is the collision diameter or size parameter and σ_{ij} is the unlike size parameter, which can be calculated by the Lorentz combining rule given by:

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \quad (10)$$

where the indices i and j refer to all the combinations between the molecules i and j .

Moreover, Hudson and McCoubrey [20] calculated the binary interaction parameter k_{ij} by equation 11 [21].

$$k_{ij} = 1 - \frac{2^7 \sqrt{I_i I_j}}{I_i + I_j} \left[\frac{\sqrt{\sigma_i \sigma_j}}{\sigma_i + \sigma_j} \right]^6, \quad 0 \leq k_{ij} \leq 1 \quad (11)$$

It can be noticed that these equations do not directly include the CPA EoS parameters. However, equation 11 can be easily converted into equation 12 if the approximate relations $\epsilon \propto a/b$ and $\sigma^3 \propto b$ are introduced [5].

$$k_{ij} = 1 - \frac{2\sqrt{I_i I_j}}{I_i + I_j} \left[\frac{\sqrt{b_i b_j}}{b_{ij}} \right]^2 \quad (12)$$

Then, a_{ij} may be calculated by equation 13 [5].

$$a_{ij} = \sqrt{a_i a_j} \frac{2\sqrt{I_i I_j}}{I_i + I_j} \left[\frac{\sqrt{b_i b_j}}{b_{ij}} \right]^2 \quad (13)$$

2.3.1 Method 1

If the Lennard-Jones 12:6 potential is not considered and a more general approach is taken using the Mie potential function, the binary interaction parameter would be calculated as:

$$k_{ij} = 1 - \frac{2\sqrt{I_i I_j}}{I_i + I_j} \left[\frac{\sqrt{b_i b_j}}{b_{ij}} \right]^{n/3-1} \quad (14)$$

where n is an adjustable exponent normally obtained from potential functions.

On the other hand, if the objective is to calculate the exponents n from known binary interaction parameters, equation 14 has to be rearranged as:

$$n = 3 \frac{\ln \left[\frac{(1-k_{ij})(I_1+I_2)}{2\sqrt{I_1 I_2}} \right]}{\ln \left[\frac{\sqrt{b_i b_j}}{b_{ij}} \right]} + 3 \quad (15)$$

2.3.2 Method 2

The previous method requires the knowledge of the ionisation potentials for all the chemical species present, which sometimes are unavailable.

Coutinho *et al.* [11, 22] developed an approximate expression for the term which involves the ionisation potentials, as the following equation shows [3]:

$$I \propto \frac{1}{b} \implies \frac{2\sqrt{I_i I_j}}{I_i + I_j} \cong \left(\frac{\sqrt{b_i b_j}}{b_{ij}} \right)^{-1} \quad (16)$$

After applying the approximation present in equation 16 in equation 14, an alternative method of determining k_{ij} is obtained:

$$k_{ij} = 1 - \left(\frac{\sqrt{b_i b_j}}{b_{ij}} \right)^{n/3-2} \quad (17)$$

The previous equation implies that the exponents n for this method are given by equation 18.

$$n = 3 \frac{\ln(1 - k_{ij})}{\ln \left[\frac{\sqrt{b_i b_j}}{b_{ij}} \right]} + 6 \quad (18)$$

3 Results

3.1 A predictive approach for the pure component parameters

3.1.1 CPA parameters database

Table 1 summarises the chemical compounds included in this work's database.

Table 1 Number of compounds per chemical family in the database compiled in this work.

Chemical Family	Number	Chemical Family	Number	Chemical Family	Number
Alcohols	16	Bromides	2	Ketones	6
Aldehydes	3	Carboxylic acids	2	Polyols	6
Alkanes	39	Chlorides	3	Sulfides	4
Alkenes	6	Cycloalkenes	3	Thiols	5
Amines	10	Esters	43	No well-defined family	7
Anhydrides	1	Ethers	8		
Aromatics	21	Fluorocarbons	12	Total	197

The CPA parameters for these 197 chemical compounds can be found from Table SA1 to SA4 and the critical properties from Table SD1 to SD4, in Supporting Information.

3.1.2 Finding trends for the CPA parameters

After having established the database, correlations were searched for the parameters. The fluorocarbons family was excluded from plot representation since they do not fit any of the equations, as already reported by Kontogeorgis and Folas [3] and Schlaikjer [10].

First, the van der Waals volume (V_w) and surface area (A_w) were predicted using the Bondi method [23], a group contribution methodology, whenever they could not be found in the literature. Next, the energy term parameters (a_0) were determined from the energy parameters (Γ) (or vice-versa), using:

$$\Gamma = \frac{a_0}{R_g b} \quad (19)$$

Then, all the parameters were plotted against V_w and their dependence was studied. First, a second-degree correlation was found for the energy term parameter (a_0) and against the van der Waals volume (V_w) for all the chemical compounds except fluorocarbons, as Fig. 1 illustrates.

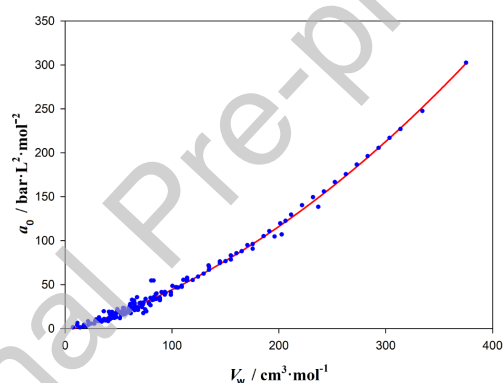


Fig 1 The energy term parameter (a_0) against the van der Waals volume (V_w) for all chemical compounds except fluorocarbons. The regression (—) follows equation $a_0 = 1.25 \cdot 10^{-3} \cdot V_w^2 + 0.34 \cdot V_w - 2.84$, with a determination coefficient of 0.9941.

Another very good correlation was found for the co-volume parameter (b) against the van der Waals volume (V_w) for all the chemical compounds except fluorocarbons, as Fig. 2 shows.

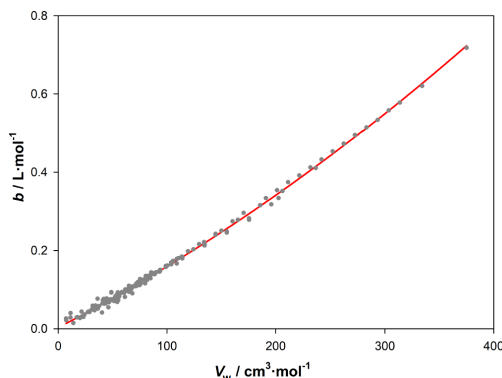


Fig 2 The co-volume parameter (b) against the van der Waals volume (V_w) for all compounds except fluorocarbons. The regression (—) follows equation $b = 1.32 \cdot 10^{-6} \cdot V_w^2 + 1.42 \cdot 10^{-3} \cdot V_w + 3.64 \cdot 10^{-3}$, with a determination coefficient of 0.9973.

On the other hand, no good correlations were found for the energy term parameter (c_1) against the van der Waals volume, as Fig. SA2 in the Supporting Information shows, so this parameter had to be fitted. Some other figures with less relevant correlations for the CPA EoS parameters can be seen in the Supporting Information.

3.1.3 Decreasing the number of CPA parameters needed

Following the establishment of reasonable trends for a_0 and b , several methods were investigated for eliminating the CPA EoS parameters for all the compounds in this work's database, with the exception of fluorocarbons, terephthalic acid and glycerol. These methods are presented in Table 2.

Table 2 Procedures taken for the CPA EoS parameters in the four different predictive methods for liquid density and vapour pressure.

Method	Energy parameter (Γ)	Co-volume (b)	Energy term parameter (c_1)	Association energy (ϵ)	Association parameter (β)
1	a	a	a	a	a
2	Correlated	Correlated	a	a	a
3	Correlated	Correlated	Fitted	a	a
4	Fitted	Correlated	Fitted	a	a

^a Literature values.

In the first method, only values from the literature were considered for the CPA EoS parameters. In the second method, correlated versions of a_0 and b were determined using the equations from Figs. 1 and 2, respectively, and the Γ were determined from the

correlated energy term parameters (a_0), using equation 19. In the third and fourth methods, some parameters were fitted using the objective function 20, while the co-volume was correlated using the equation from Fig. 2.

$$\text{OF} = \sum_{j=1}^{N_D} \left[\left(\frac{P_{\text{CPA},i,j}^{\text{sat}} - P_{\text{exp},i,j}^{\text{sat}}}{P_{\text{exp},i,j}^{\text{sat}}} \right)^2 + \left(\frac{\rho_{\text{CPA},i,j}^{\text{sat}} - \rho_{\text{exp},i,j}^{\text{sat}}}{\rho_{\text{exp},i,j}^{\text{sat}}} \right)^2 \right] \quad (20)$$

where N_D denotes the number of data points, $P_{\text{CPA},i,j}$ is the saturation pressure predicted by the CPA EoS model, $P_{\text{exp},i,j}^{\text{sat}}$ is the experimental saturation pressure, $\rho_{\text{CPA},i,j}$ is the saturated liquid density predicted by the CPA EoS model and $\rho_{\text{exp},i,j}^{\text{sat}}$ is the experimental saturated liquid density.

In this objective function, a temperature range of $0.50 \cdot T_c$ to $0.95 \cdot T_c$ and a step of 1 K were considered. The upper limit of the critical temperature interval follows the works of Michelsen [24]. In the Supporting Information, Appendix D, the critical properties for all chemical compounds except for fluorocarbons can be seen. When not found in the literature, the critical temperatures were predicted using the Joback method [25].

Furthermore, no experimental data were collected, but correlations from the Design Institute for Physical Properties (DIPPR) [26] and based on experimental data were used.

For all the methods, the average absolute deviation (AAD) was calculated following equation 21 and considering the data obtained from the DIPPR correlations [26] as reference.

$$\text{AAD} = \sum_{i=1}^{N_C} (\text{AD}_i) / N_C \quad (21)$$

where AD_i is the average absolute deviation of component i , N_C stands for the total number of components and AD stands for the absolute relative deviation, calculated by equation:

$$\text{AD}_i = \sum_{j=1}^{N_D} \left(\left| \frac{Y_{\text{method},j} - Y_{\text{DIPPR},j}}{Y_{\text{DIPPR},j}} \right| \right) / N_D \cdot 100 \% \quad (22)$$

where $Y_{\text{method},j}$ is a generic property calculated by one of the methods for data point number j and $Y_{\text{DIPPR},j}$ is the value of that property when following the DIPPR correlations [26]. N_D stands for the total number of data points.

Although, some chemical species were not present in the DIPPR database, which caused a decrease to 174 in the total number of analysed chemical compounds. Table

3 summarises the obtained results. The results for each chemical species and all the correlated and fitted parameters can be seen in the Supporting Information, from Table SA5 to SA9 and from Table SA10 to SA13, respectively.

Table 3 Average absolute deviations (AAD) and convergence rate for the predictions of saturated vapour pressure and saturated liquid density per method.

Method	Vapour pressure	Liquid density	Convergences / %
	AAD / %	AAD / %	
1	3.7	2.0	100
2	59.4	6.0	99.0
3	14.1	4.7	99.5
4	1.5	3.7	100

From Table 3, it is obvious that the AAD are always smaller in the first method, when compared to the second and third methods. This fact was expected since the CPA EoS parameters are obtained by fitting their value to experimental data. Fitting the energy term parameter (c_1) in method 3 significantly reduced the deviations and increased the number of convergences, even though significant deviations were still observed for the vapour pressures.

Moreover, the fitting of the energy term parameter (Γ) ensured the same number of convergences observed in the first method and provided very similar results, being slightly more accurate describing vapour pressures and less precise predicting liquid densities.

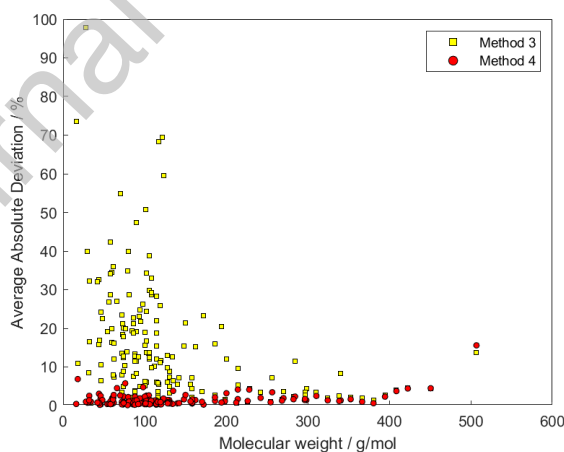
For non-associating compounds, the third and fourth methods only used either correlated or fitted parameters, but still obtained relatively small deviations, which are very important results towards an increased predictive ability.

In Table 4, the results can be seen for each considered chemical family.

Table 4 Average absolute deviations (AAD) for the liquid densities and vapour pressure per chemical family.

Chemical family	Nr. of compounds	Saturated liquid density				Saturated vapour pressure			
		Methods				Methods			
		1	2	3	4	1	2	3	4
		AAD / %	AAD / %	AAD / %	AAD / %	AAD / %	AAD / %	AAD / %	AAD / %
Alcohols	16	1.1	2.2	2.5	3.0	2.7	31.9	10.3	1.3
Alkanes	39	2.3	5.2	3.7	2.4	4.6	36.2	12.4	1.8
Alkenes	6	0.8	2.9	2.1	1.8	2.5	36.7	13.0	1.1
Amines	10	3.4	6.6	5.2	4.2	4.4	70.2	20.8	0.7
Aromatics	20	0.8	8.5	4.6	2.5	1.2	170.5	23.7	0.9
Esters	43	2.3	3.5	3.1	2.8	4.4	29.7	6.2	1.4
Ethers	8	0.9	7.0	4.9	3.0	1.3	57.3	16.6	1.0
Ketones	6	0.8	2.9	2.1	1.8	2.5	36.9	13.0	1.1
Polyols	5	4.1	9.0	7.9	8.1	9.5	108.3	19.3	1.7
Thiols	5	2.1	5.0	3.5	2.1	7.1	30.3	13.2	2.1
Associating	45	1.3	7.0	5.9	5.5	3.5	73.5	15.7	1.3
Non-associating	129	1.9	5.3	3.8	2.8	3.8	60.1	12.9	1.4
All	174	2.0	6.0	4.7	3.7	3.7	59.4	14.1	1.5

The AAD in vapour pressure for method 3 and 4 were plotted against the molecular weight and the results can be seen in Fig. 3. It can be observed that the fitting of the energy parameter (Γ) in method 4 significantly reduced the deviations in the prediction of the vapour pressure.

**Fig 3** Average absolute deviations (AAD) in the calculation of vapour pressure following methods 3 and 4 against the molecular weight.

The AAD for the liquid density in these two methods were also plotted against the molecular weight, as Fig. 4 illustrates. This figure shows that the deviations remained almost unaffected for the liquid densities from method 3 to method 4, which implies that the determination of the vapour pressures is more sensitive to variations in Γ than the determination of liquid densities.

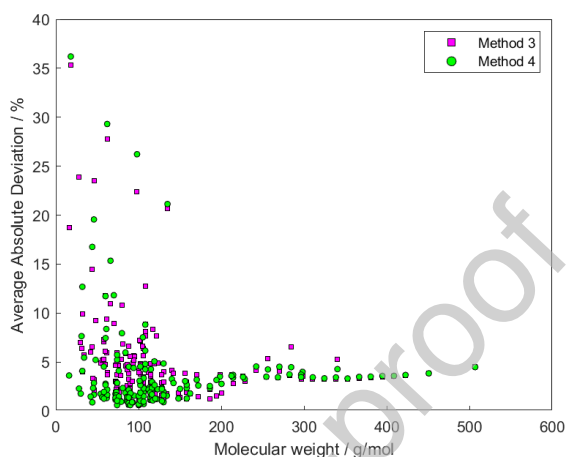


Fig 4 Average Absolute Deviations (AAD) in the calculation of liquid density following methods 3 and 4 against the molecular weight.

3.2 A predictive approach for the binary interaction parameters

3.2.1 Finding trends for the ionisation potentials

As previously seen, the ionisation potentials are normally required in the combining rules, so correlations involving them could be very useful. However, only for alkanes a good correlation was found for the ionisation potentials against the van der Waals volume, as Fig. 5 shows.

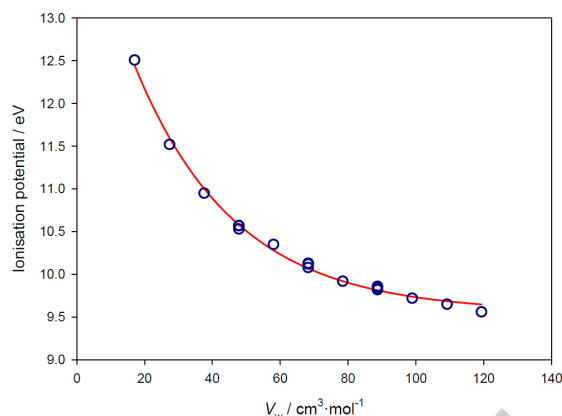


Fig 5 The ionisation potential, I , against the van der Waals volume, V_w , in alkanes. The regression (–) follows equation $I = 9.57 + 5.14 \cdot e^{-3.38 \cdot 10^{-2} \cdot V_w}$, with a determination coefficient of 0.9958.

For alcohols and ethers, a general decrease of the ionisation potentials was observed with increasing van der Waals volume, but the number of chemical compounds was insufficient to establish a satisfactory correlation.

Regarding previous efforts on this topic, Haslam *et al.* [27] showed a decreasing tendency of the ionisation potential with the carbon number in n-alkanes and Coutinho *et al.* [11] showed an increasing tendency of the ionisation potential with the inverse of the cubic molecular diameter for some hydrocarbons and gases. Although the same decreasing behaviour was found in this work, an exponential decay was observed and not an approximate linear correlation as Coutinho *et al.* [11] described. This fact is justified by the larger set of compounds studied in this work.

3.2.2 Finding trends for the binary interaction parameters

First, the methods summarised in Table 5 were used to calculate the binary interaction parameters (k_{ij}). Their theoretical and calculated values can be observed in the Supporting Information, from Table SB1 to SB8 and from Table SB9 to SB13, respectively.

Table 5 Methods used to calculate the binary interaction parameters (k_{ij}) and their respective exponents n .

Method	Comment	Use of ionisation potentials	Calc. of k_{ij}	Calc. of n
1	Hudson-McCoubrey theory [20] fitting n	Yes	eq. 14	eq. 15
2	Coutinho <i>et al.</i> [11, 22] approximation	No	eq. 17	eq. 18

Then, correlations were searched for the exponents n of the Mie potential function and for the binary interaction parameters (k_{ij}) with measures of system size asymmetries, for

both the methods mentioned. Specifically, the system asymmetries studied were ratios of size parameters, as, for example, the ratio of van der Waals volumes and the ratio of co-volumes. In this work, the ratios were defined as "larger ratios", i.e., with the larger value in the numerator and the smaller in the denominator.

For alkanes-water, good relations were obtained for the calculated exponent n with the larger co-volumes ratio for both methods, as Fig. 6 shows.

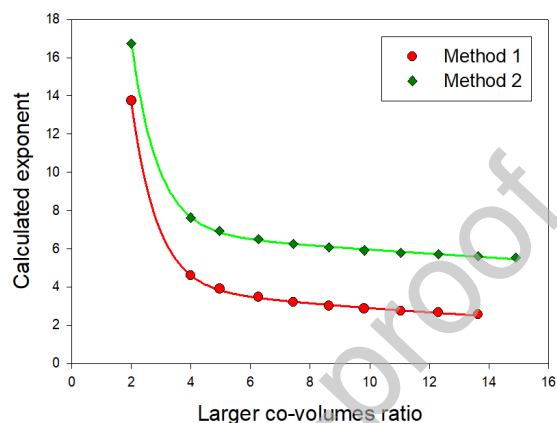


Fig 6 Calculated exponent n against the larger ratio of co-volumes in alkanes-water for both methods. The regression for the first method (–) follows equation $n = 1.04 \cdot 10^2 \cdot e^{-1.18 \cdot b_{\text{ratio}}} + 4.32 \cdot e^{-4.01 \cdot 10^{-2} \cdot b_{\text{ratio}}}$, with a determination coefficient of 0.9998. The regression for the second method (–) follows equation $n = 9.58 \cdot 10^1 \cdot e^{-1.13 \cdot b_{\text{ratio}}} + 7.12 \cdot e^{-1.79 \cdot 10^{-2} \cdot b_{\text{ratio}}}$, with a determination coefficient of 0.9997.

For acetic acid-alkanes, some good correlations were found for the calculated exponent n with the van der Waals volumes ratio and with the co-volumes ratio for both methods. In Fig. 7 the relation between the exponents n and the ratio of co-volumes can be seen for the second method.

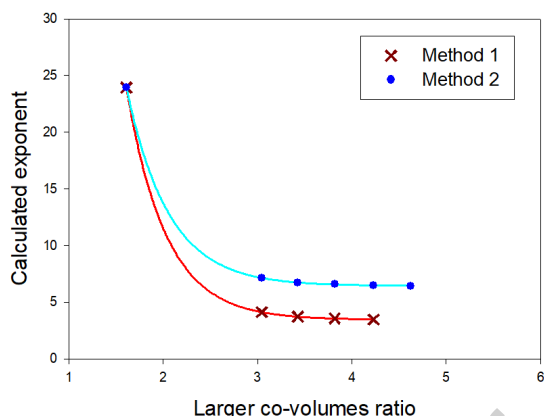


Fig 7 Calculated exponent n against the ratio of co-volumes in acetic acid-alkanes for both methods. The regression for the first method (–) follows equation $n = 3.46 + 9.52 \cdot 10^2 \cdot e^{-2.39 \cdot b_{\text{ratio}}}$, with a determination coefficient of 1.0000. The regression for the second method (–) follows equation $n = 6.45 + 6.44 \cdot 10^2 \cdot e^{-2.24 \cdot b_{\text{ratio}}}$, with a determination coefficient of 1.0000.

Table 6 summarises the majority of the results for the second method, which is the most practical and the one with the largest number of data points.

Table 6 Summary of the relations found for the exponents n and for the binary interaction parameter (k_{ij}) with the larger van der Waals volumes ratio (V_{ratio}) and with the larger co-volumes ratio (b_{ratio}) in the second method for different chemical pairs.

Chemical pairs	Exponent n		k_{ij}	
	V_{ratio}	b_{ratio}	V_{ratio}	b_{ratio}
Acetic acid-Alkanes	✓	✓	×	×
Acetone-Hydrocarbons	✓	✓	✓	✓
Alkanes-Ethanol	✓	✓	×	×
Alkanes-Ethers	×	×	×	×
Alkanes-Glycols	✓	✓	×	×
Alkanes-MEG	×	✓	×	×
Alkanes-Polyols	✓	✓	×	×
Alkanes-Water	✓	✓	✓	✓
Aromatics-Water	×	×	×	×
BTEX-Benzene	✓	✓	✓	✓
Ethers-Water	×	×	×	×

✓ - Good relation found. × - No good relation found.

As Table 6 shows, it was easier to find good correlations for the exponent n than for the binary interaction parameter. Overall, reasonable values for the exponents (between 0 and 10) were obtained using both methods applied.

Though, when the species in the chemical pair under study were too similar (small size asymmetries), the denominator in equation 18 became almost zero and the exponents (n) assumed high non-physical values, as Fig. 8 shows for the second method. For ratios of co-volumes higher than 2, n resulted to reasonable values. Moreover, if the ratio became even higher, an almost constant value of the exponents was achieved.

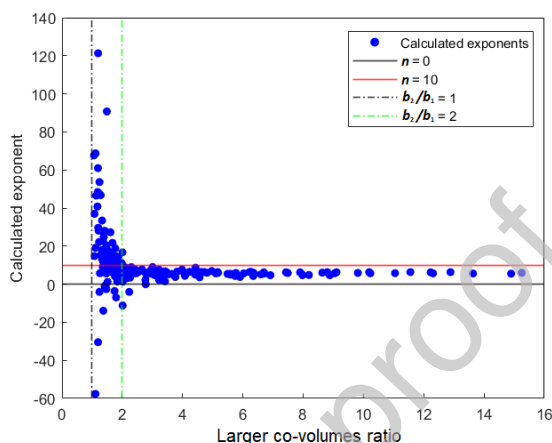


Fig 8 Calculated exponents n against the larger co-volumes ratio for all the combinations with the second method.

3.2.3 Sensitivity analysis

In order to study the sensitivity of the second method to the use of correlations, the exponents of the Mie potential function were correlated for two different systems (alkanes-water and acetic-acid alkanes), using the correlations from Figs. 6 and 7, respectively. Then, the k_{ij} were calculated with the correlated exponents using equation 17.

Fig. 9 shows both the predicted and theoretical binary interaction parameters against the larger co-volumes ratio for alkanes-water.

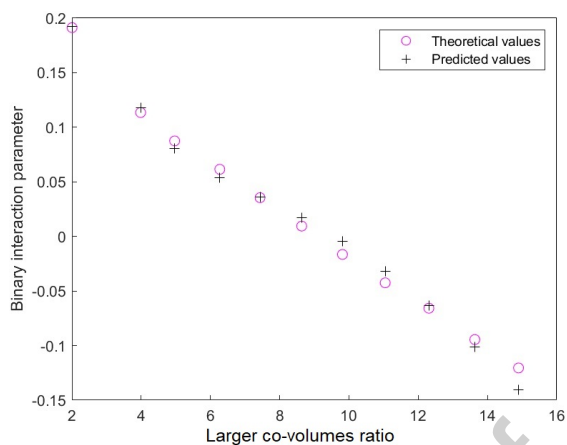


Fig 9 Predicted and theoretical binary interaction parameters, k_{ij} , in alkanes-water for the second method.

Fig. 10 shows both the predicted and theoretical binary interaction parameters against the larger co-volumes ratio for alkanes-water.

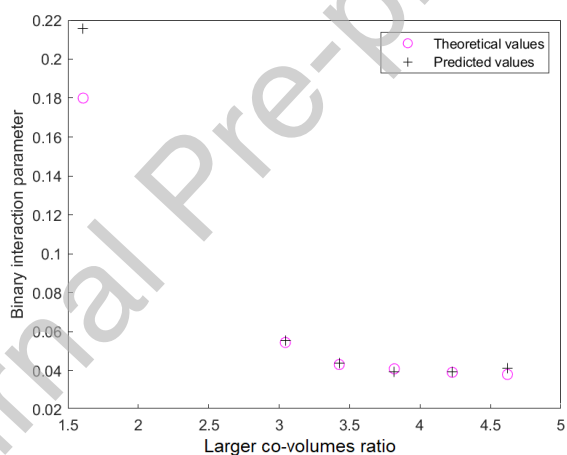


Fig 10 Predicted and theoretical binary interaction parameters, k_{ij} , in acetic acid-alkanes for the second method.

In both cases, the correlated exponents for the second method were able to predict the binary interaction parameters (k_{ij}) without requiring the knowledge of the ionisation potentials.

So as to study the sensitivity in a more general perspective, different positive and negative variations in the exponents n and their respective influences on the values of the binary interaction parameters (k_{ij}) were analysed. Fig. 11 shows the comparison between

the theoretical binary interaction parameters (from the database) and the predicted ones (affected by variations in the exponents n).

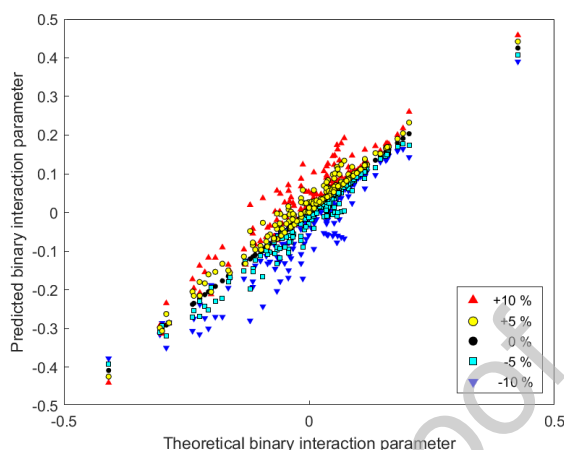


Fig 11 Effect of variations of the exponents n in the determination of the binary interaction parameters (k_{ij}) using the second method.

As it can be seen from Fig. 11, the second method is very sensitive to small variations in the values of the exponents, and this fact may hamper the usage of correlations in some chemical families. However, Fig. 11 does not provide any information on how these variations in the value of the binary interaction parameters would affect the prediction of binary data. Some systems are very sensitive to k_{ij} while others are not.

Hence, to compare the performances of the literature-based binary interaction parameters with the correlated (using equation (17) and regressed ones (using general trends), two extra plots are provided in the Supporting Information, Appendix E.

For n-heptane-water, the description of temperature against the mutual solubility (liquid-liquid equilibrium T_{xy} diagram) using the literature-based binary interaction parameter was compared to the obtained using the correlated k_{ij} , calculated using equation (17), and to the determined using the regressed value. The regressed value of k_{ij} was obtained from the equation for the second method shown in Fig. 6. The results can be observed in Fig. SE1.

The same comparison was performed for acetic acid-n-octane in the representation of pressure against the mole fraction of acetic acid (P_{xy} diagram) at 400 K, which can be seen in Fig. SE2. The calculated k_{ij} was determined using equation (17) and the regressed one was obtained from the equation for the second method presented in Fig. 7.

The three methods of obtaining the binary interaction parameters presented very similar results, with the correlated parameters describing better the studied properties than the

regressed ones. These approximate methods resulted to reasonable values of k_{ij} for the studied systems, thus validating the approach.

4 Discussion

4.1 Pure component parameters

4.1.1 A predictive approach for the pure component parameters

It was possible to estimate both the energy term parameter (a_0) and the co-volume (b) based only on the van der Waals volume, which can be predicted using a group contribution method as the Bondi [23] method. Even though this is a very important step towards a more predictive CPA EoS, no correlations were found for the c_1 parameter nor for the association parameters (ϵ and β), which still need to be fitted to experimental data.

Although the CPA EoS parameters database (see Supporting Information) comprises a relatively large number of chemical compounds, the species with van der Waals volumes lower than $100 \text{ cm}^3 \cdot \text{mol}^{-1}$ are far more abundant than the others and they usually include the compounds with the highest deviations from the correlations.

On the other hand, in the high-values zone of V_w , only long-chain alkanes are present, which are well-predicted by the fits. Therefore, in order to better understand the influence of the high molecular mass alkanes in these correlations, alternative representations were performed, in which the species with molecular mass higher than $240 \text{ g} \cdot \text{mol}^{-1}$ were removed. This way, a decrease of the determination coefficients for the same type of correlations was observed from 0.9941 to 0.9758 in Fig. 1 and from 0.9973 to 0.9918 in Fig. 2. As expected, this part of the alkanes' family improved all the results when taken into account. However, even without considering them, the correlations are still considered very useful.

The fluorocarbons family presented a different behaviour than the other chemical families. However, no explanation could be found, so this chemical family was not considered in parameter estimation. When the fluorocarbons were taken into account, the determination coefficient of the correlations was significantly lower, as it can be seen from Fig. SA3 to SA6 in the Supporting Information. The chemical compound which deviated the most from the common trends was perfluorodecalin.

When comparing the van der Waals surface area (A_w) with the van der Waals volume (V_w), it was noticed that the van der Waals volume generally provided correlations with higher determination coefficients, so only these were kept. This difference was minimal for alkanes, but critical for associating species. Regarding the energy term parameters a_0 and Γ , it was observed that the latter usually provides correlations with smaller determination coefficients and that needs more complex functions with a higher number of adjustment parameters. As a result, the energy term parameter a_0 seems more suitable as a correlating

parameter and should be preferred to the energy term parameter Γ . In the Supporting Information, some tables make further comparisons between the mentioned parameters and properties.

4.1.2 *Decreasing the number of CPA parameters needed*

In Table 4, it can be seen that the obtained results were very close to method 1 (theoretical values), which is a very important achievement, since no parameters from the literature were used in non-associating compounds for methods 3 and 4. For the vapour pressures obtained with method 4, the results were even better than with the first method, however, since the CPA parameters are fitted to experimental data, an approximate method like method 4 should not present better results than method 1, so this gap is probably due to choosing different reference databases for the experimental data and/or differences in the considered values of chemical properties, for example, in the critical temperatures.

When a comparison between the energy term parameter a_0 , the co-volume (b) and the energy term parameter Γ) and their correlated values was established, it was noticed that the first two parameters presented high determination coefficients for a first-degree regression, i.e. their values and their correlated ones were almost the same, specifically, with 0.9943 and 0.9977. Although, when it came to the energy parameter (Γ) the determination coefficient drastically fell to 0.7887, which means that the correlated Γ is, in general, very different from the theoretical value, as can be seen in Fig. 12. This way, it seemed obvious that the prediction of the energy parameter (Γ) is heavily affected by small deviations of the co-volume and of the energy term parameter (a_0). Being so, this disparity may explain some deviations in the second and third methods, which were diminished by the fitting of Γ in the fourth method.

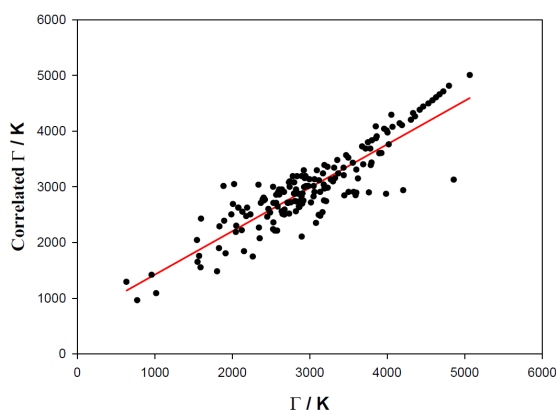


Fig 12 Comparison between the energy parameter values stored in the database of this work, Γ , and the correlated ones, $\Gamma^{\text{corr.}}$. The regression (–) follows equation $\Gamma^{\text{corr.}} = 0.80 \cdot \Gamma + 642.53$, with a determination coefficient of 0.7887.

If Fig. 3 is observed, a general decrease of the value of deviations can be observed for the vapour pressure in the method 3 with the molecular mass, which is a good result since the majority of the compounds with unknown parameters, so for which a more predictive model is needed, are heavy compounds. In Fig. 4, it can also be seen that the deviations in the liquid density almost haven't changed from method 3 to method 4 even though Γ was fitted, showing once more that the calculations of vapour pressure are significantly more sensitive to variations in Γ than the ones of liquid density.

Finally, it can be said that it was possible to find trends between the parameters of the CPA EoS and properties (or other parameters) and to successfully use these correlations to predict the values of the CPA parameters whenever sufficient experimental data are not available. Very promising results were obtained and it was possible to decrease the number of required parameters from the literature, contributing to an increased predictive ability of the CPA EoS.

4.2 A predictive approach for the binary interaction parameters

4.2.1 Finding trends for the parameters

The second method (Coutinho *et al.* [11, 22] approximation) successfully described the behaviour of the analysed systems in a very straightforward way, since it does not require the knowledge of the ionisation potentials. For very asymmetrical combinations, the exponents became almost constant, as Fig. 8 shows, which may be very useful to model mixtures with very asymmetrical compounds. However, as seen in Fig. 11, it is a very sensitive model to variations in the value of the exponents, which may hamper its usage in some systems. Moreover, when the species in the chemical pair under study were too

similar (small size asymmetries), the exponents assumed high and nonphysical values, as seen in Fig. 8. Hence, although this method may be very useful in asymmetrical mixtures in which no sufficient data exist, the usage of the Hudson-McCoubrey theory [20] with a fitted exponent (first method) should be preferred in more symmetrical systems (with co-volumes ratios smaller than 2), since it is the only model capable of accounting for the energetic differences between very similar chemical compounds.

4.3 List of correlations

The following table summarises the best correlations found, their coefficient of determination and the chemical species for which they are valid.

Table 7 Summary of the best correlations found, their respective coefficient of determination, chemical compounds for which are valid and respective figure.

Mathematical expression ^a	R^2	Species	Figure
$b = 1.32 \cdot 10^{-6} \cdot V_w^2 + 1.42 \cdot 10^{-3} \cdot V_w + 3.64 \cdot 10^{-3}$	0.9973	b	2
$a_0 = 1.25 \cdot 10^{-3} \cdot V_w^2 + 0.34 \cdot V_w - 2.84$	0.9941	b	1
$b = -2.58 \cdot 10^{-6} \cdot (T_c/P_c)^2 + 3.02 \cdot 10^{-3} \cdot T_c/P_c + 0.03$	0.9938	b	SA7
$b = -2.53 \cdot 10^{-6} \cdot a_0^2 + 3.01 \cdot 10^{-3} \cdot a_0 + 0.03$	0.9925	b	SA1
$b = 4.66 \cdot 10^{-7} \cdot MW^2 + 1.17 \cdot 10^{-3} \cdot MW + 3.47 \cdot 10^{-3}$	0.9996	Alkanes	SA9
$a_0 = 7.13 + 0.299 \cdot e^{0.00783 \cdot T_c}$	0.9981	Alkanes	SA12
$I = 9.57 + 5.14 \cdot e^{-3.38 \cdot 10^{-2} \cdot V_w}$	0.9958	Alkanes	5
$c_1 = 0.228 + 0.136 \cdot e^{0.00297 \cdot T_c}$	0.9954	Alkanes	SA14

^a b is in L/mol, V_w in cm^3/mol , a_0 in $\text{bar} \cdot \text{L}^2 \cdot \text{mol}^{-2}$, T_c in K, P_c in MPa, MW in g/mol and I in eV.

^bAll except fluorocarbons.

5 Conclusions

For a database with CPA EoS parameters and properties for 197 chemical compounds, correlations between parameters and properties were searched for all compounds and between groups of similar species. Using the correlations found for pure compounds, four different methods were used in predicting vapour pressure and saturated liquid density. The predicted densities and pressures were compared to the correlations from DIPPR and, in all the methods, a temperature range from $0.5 \cdot T_c$ to $0.95 \cdot T_c$ with a step of 1 K was considered. It was possible to obtain small deviations using correlated and fitted CPA parameters.

In a database with 219 binary interaction parameters (k_{ij}), the Mie potential function exponents (n) were calculated using the Hudson-McCoubrey theory [20] and following two different methods. In general, very reasonable values were obtained for the exponents (between 0 and 10). The second method, an approximation developed by Coutinho *et al.* [11,22], ended up being a very good alternative to the classical theory, correctly describing the species behaviours even though it had a simplified mathematical expression. Though, the model was considered sensitive to variations in the exponents n and that this fact could hamper the future usage of correlations in some chemical families. Moreover, when the species in the chemical pair under study were too similar (small size asymmetries), the exponents assumed high and nonphysical values. Hence, although this method may be very useful for asymmetrical mixtures in which no sufficient data exist, the usage of the Hudson-McCoubrey theory [20] with a fitted exponent (first method) should be preferred in more symmetrical systems, since it is the only model capable of accounting for the energetic differences between very similar chemical compounds.

In conclusion, by exploring the common behaviour of the properties of similar compounds, it was possible to generalise some correlations and to use them in the determination of vapour pressure, liquid density and binary interaction parameters, successfully increasing the predictive ability of the CPA EoS.

Acknowledgements

This work was financially supported by Base Funding - UIDB/50020/2020 of the Associate Laboratory LSRE-LCM - funded by national funds through FCT/MCTES (PID-DAC). On the other hand, Eléna Gmez thanks funding support from Stimulus of Scientific Employment, Individual Support 2017, CEECIND/02646/2017.

Supporting Information Introduction

In the Supporting Information, the compiled databases and the predicted parameters and properties can be seen. Additionally, some figures present extra correlations and others illustrate the atypical behaviour of the fluorocarbons' family.

References

- 1 G. M. Kontogeorgis, E. C. Voutsas, I. V. Yakoumis, and D. P. Tassios. An equation of state for associating fluids. *Ind. Eng. Chem. Res.*, 35:4310–4318, 1996.
- 2 G. Soave. Equilibrium constants from a modified Redlich-Kwong equation of state. *Chem. Eng. Sci.*, 27:1197–1203, 1972.
- 3 G. M. Kontogeorgis and G. K. Folas. *Thermodynamic Models for Industrial Applications: From Classical and Advanced Mixing Rules to Association Theories*. John Wiley & Sons, 1st edition, 2010.

- 4 G. M. Kontogeorgis. Association theories for complex thermodynamics. *Chem. Eng. Res. Des.*, 91:1840–1858, 2013.
- 5 T. A. Hansen. Development of a predictive version of the CPA equation of state for applications in the chemical industry. Master's thesis, Technical University of Denmark, Denmark, 2015.
- 6 D. W. van Krevelen and K. te Nijenhuis. *Properties of Polymers*, page 73. Elsevier Science, 4th edition, 2009.
- 7 A. M. Palma, M. B. Oliveira, A. J. Queimada, and J. A. P. Coutinho. Evaluating cubic plus association equation of state predictive capacities: A study on the transferability of the hydroxyl group associative parameters. *Ind. Eng. Chem. Res.*, 56:7086–7099, 2017.
- 8 P. Mathias and T. Copeman. Extension of the peng-robinson equation of state to complex mixtures: evaluation of the various forms of the local composition concept. *Fluid Phase Equilib.*, 13:91–108, 1983.
- 9 S. Ali. Application of the CPA EoS to CO₂-containing systems. Master's thesis, Technical University of Denmark, Denmark, 2011.
- 10 A. Schlaikjer. Towards a predictive CPA equation of state for chemical applications. Master's thesis, Technical University of Denmark, Denmark, 2014.
- 11 J. A. P. Coutinho, P. M. Vlamos, and G. M. Kontogeorgis. General form of the cross-energy parameter of equations of state. *Ind. Eng. Chem. Res.*, 39:3076–3082, 2000.
- 12 M. L. Michelsen and E. M. Hendriks. Physical properties from association models. *AIChE Journal*, 180:165–174, 2001.
- 13 G. M. Kontogeorgis, I. V. Yakoumis, H. Meijer, E. Hendriks, and T. Moorwood. Multicomponent phase equilibrium calculations for water-methanol-alkane mixtures. *Fluid Phase Equilib.*, 158-160:201–209, 2001.
- 14 J. R. Elliott Jr., S. J. Suresh, and M. D. Donohue. A simple equation of state for non-spherical and associating molecules. *Ind. Eng. Chem. Res.*, 29:1476–1485, 1990.
- 15 G. K. Folas, J. Gabrielsen, M. L. Michelsen, E. H. Stenby, and G. M. Kontogeorgis. Application of the Cubic-Plus-Association (CPA) equation of state to cross-associating systems. *Ind. Eng. Chem. Res.*, 44:3823–3833, 2005.
- 16 S. H. Huang and M. Radosz. Equation of state for small, large, polydisperse, and associating molecules. *Ind. Eng. Chem. Res.*, 29:2284–2294, 1990.
- 17 V. S. Gangadhara Rao and M. Mukhopadhyay. Solid solubilities in supercritical fluids from group contributions. *J. Supercrit. Fluids*, 3:66–70, 1990.
- 18 U. Plcker, H. Knapp, and J. Prausnitz. Calculation of high-pressure vapor-liquid equilibria from a corresponding-states correlation with emphasis on asymmetric mixtures. *Ind. Eng. Chem. Process. Des. Dev.*, 17:324–332, 1978.
- 19 T. M. Reed III. The theoretical energies of mixing for fluorocarbon-hydrocarbon mixtures. *J. Chem. Phys.*, 59:425–428, 1955.
- 20 G. H. Hudson and J. C. McCoubrey. Intermolecular forces between unlike molecules: a more complete form of the combining rules. *Trans. Faraday Soc.*, 56:761–771, 1960.
- 21 M. Frost, G. K. Folas, N. Muro-Sune, F. R. Leon, M. P. Breil, N. von Solms, M. L. Michelsen, and G. M. Kontogeorgis. The Cubic-Plus-Association EoS: Parameters for pure compounds and interaction parameters. Technical report, Technical University of Denmark, September 2012.

- 22 J. A. P. Coutinho, G. M. Kontogeorgis, and E. H. Stenby. Binary interaction parameters for nonpolar systems with cubic equations of state: a theoretical approach 1. CO₂/hydrocarbons using SRK equation of state. *Fluid Phase Equilib.*, 102:31–60, 1994.
- 23 A. Bondi. Van der Waals volumes and radii. *J. Phys. Chem.*, 68:441–451, 1964.
- 24 M. L. Michelsen. Parameter estimation in association models. Proceedings of ESAT, France 2008.
- 25 K. G. Joback and R. C. Reid. Estimation of pure-component properties from group-contributions. *Chem. Eng. Commun.*, 57:233–243, 1987.
- 26 Design Institute for Physical Properties - AIChE. DIADEM Professional Data Evaluation Manager. Version 12.3.0, May 2018.
- 27 A. J. Haslam, A. Galindo, and G. Jackson. Prediction of binary intermolecular potential parameters for use in modelling fluid mixtures. *Fluid Phase Equilib.*, 266:105–128, 2008.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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

This work was financially supported by Base Funding - UIDB/50020/2020 of the Associate Laboratory LSRE-LCM - funded by national funds through FCT/MCTES (PIDDAC).

On the other hand, Elena Gómez acknowledges funding support from Stimulus of Scientific Employment, Individual Support 2017, CEECIND/02646/2017.