Predicting solvation energies of free radicals and their mixtures: A robust approach coupling the Peng-Robinson and COSMO-RS models

Francisco Carlos Paes, Romain Privat, Jean-Noël Jaubert, Baptiste Sirjean

Predicting Solvation Energies of Free Radicals and Their Mixtures: A Robust Approach Coupling the Peng-Robinson and COSMO-RS models

Francisco Paes[1], Romain Privat[1(*)], Jean-Noël Jaubert[1(*)], Baptiste Sirjean[1(*)]

[1] *Université de Lorraine, CNRS, LRGP, 1 rue Grandville, 54000 Nancy, France*

(*) Corresponding authors: romain.privat@univ-lorraine.fr
jean-noel.jaubert@univ-lorraine.fr
baptiste.sirjean@univ-lorraine.fr

Abstract

Detailed kinetic models are valuable tools for a clear understanding of the dynamics of complex reacting systems based on radical-chain mechanisms, such as combustion and oxidation. Gas-phase automatic kinetic generators, such as EXGAS, RMG, MAMOX, NetGen, REACTION, and GENESYS can be used to establish a detailed network of reaction pathways and the rates associated with each reaction. The range of applications of these gas-phase kinetic generators can be extended to liquids through diffusional and solvation corrections directly applied to the gas-phase thermo-kinetic data. In this work, a flexible framework is proposed for the calculation of the solvation correction involving closed-shell (i.e., non-free radical) and open-shell (i.e., free radicals) molecules. This novel model relies on the Peng-Robinson cubic equation of state (EoS) combined with a quantum-based continuum solvation model (COSMO-RS) through an advanced mixing rule. Unlike most predictive equations of state, the proposed model requires only pure compound inputs: critical temperature ($T_{c,i}$), critical pressure ($P_{c,i}$), acentric factor (ω_i), and the screening charge distribution (σ-profile). These inputs were obtained for C/H/O free radicals using group corrections applied to the known values of the associated closed-shell molecules (parent molecules). The resulting EoS is able to provide fast predictions of solvation quantities of closed-shell molecules and C/H/O free radicals with a mean unsigned error of around 0.30 kcal/mol. The great advantage of this method is that it allows high-throughput computation of solvation quantities in pure solvents and mixtures at any temperature (including the supercritical domain), which could enable the simulation of complex oxidation kinetics in a wide range of applications.

Keywords: liquid phase kinetics; solvation Gibbs energy; free radicals; cubic equation of state; COSMO-RS

1 Introduction

Kinetic models offer valuable insights into the dynamic behavior of reacting systems. Constructing a kinetic model involves creating a network of chemical reactions that accurately depict the underlying mechanism of a given chemical process. Subsequently, each reaction is attributed a rate law, shedding light on the temporal progression of these reactions.

For simple systems involving a small number of reactions, the construction of the mechanism and its pathways can be done manually based on the knowledge acquired through other disciplines such as organic chemistry and biochemistry [1]. However, when dealing with complex chemical processes characterized by extensive networks of reactions, the task of manually constructing a kinetic model can quickly become too laborious or even unfeasible. Fortunately, recent strides in computational chemistry have opened the door to the implementation of the so-called automatic kinetic generators, which alleviate this challenge by enabling the automated construction of complex reaction networks.

In short, automatic kinetic generators are based on perfect gas approximations for the computation of the required thermochemical properties (often employing the group additivity method originally formulated by Benson [2]), along with established rate rules and predefined reaction templates [3]. Some examples of automatic kinetic generators are EXGAS, RMG, MAMOX, NetGen, REACTION, and GENESYS. They have been widely used to describe complex free radical chain mechanisms taking place in gas phase, such as pyrolysis, combustion or atmospheric oxidation phenomena [4].

Nonetheless, in some applications of industrial interest, those oxidation mechanisms take place in liquid phase. A notable illustration of this is the Hock process, where cumene undergoes oxidation reactions in a liquid medium to yield phenol [5,6]. This process attracts a lot of attention because phenol is a crucial precursor for the synthesis of a wide range of organic chemicals and polymers. Another noteworthy example is the oxidation of cyclohexane to produce KA oil (a mixture of cyclohexanone K and cyclohexanol A), which is a vital intermediate in the Nylon manufacturing process [7,8]. The liquid-phase oxidation mechanism can also play an important role in advanced wastewater treatments. An intriguing illustration can be found in the Fenton oxidation process [9], where hydroxyl free radicals are generated through the reaction of hydrogen peroxide with an iron catalyst to promote the degradation of organic pollutants. Regarding the energetic transition context, liquid phase oxidation is also the driven mechanism behind fuel ageing, which turns to be a major drawback when it comes to the use of biofuels in aircraft [10–12]. Modelling the physical process governing fuel ageing is therefore a necessary first step to control and limit this undesirable phenomenon.

The above examples demonstrate that there is a demand for kinetic models capable of describing complex mechanisms in liquid phase. However, and in agreement with the literature, developing detailed kinetic models for such cases remains a challenge mainly because of the thousands of solvent-dependent thermo-kinetic data they demand. Yet, not only are thermo-kinetic experimental data scarcer in pure solvents, and virtually non-existent for complex mixtures, but there are no liquid-phase data for the free radicals at the heart of chain radical mechanisms. Because of this, the approach adopted in the literature is to extend the gas-phase mechanism to the condensed phase by adding corrections to the gas-phase thermochemical and kinetic data. For instance, rate constants for bimolecular reactions can be extended to liquids by incorporating diffusional and solvation effects, as shown in equation (1) [13].

$$k_{liq}^{eff} = \frac{4\pi RDk_{liq}^{int}}{4\pi RD + k_{liq}^{int}} \quad (1)$$

Where, k_{liq}^{eff} is the effective reaction rate in liquid-phase, the term $4\pi RD$ is the rate of diffusion, in which R is the sum of radii of the reactants, and D is the sum of the diffusivities of the reacting species that can be estimated by the Stokes-Einstein relation [14]. The intrinsic reaction rate k_{liq}^{int} is obtained by coupling solvent-solute interactions with the gas phase kinetic constant (k_{gas}), as follows:

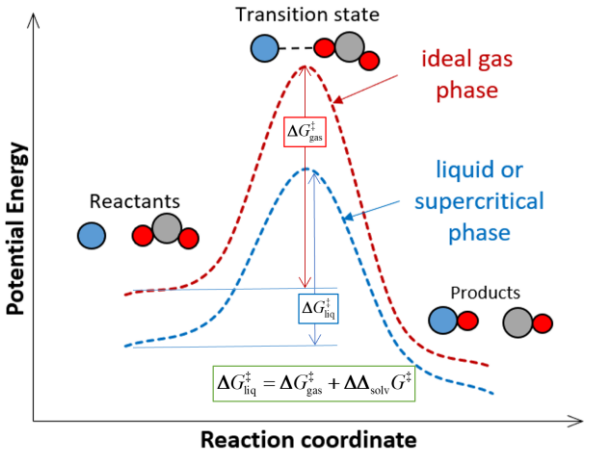
$$k_{liq}^{int} = k_{gas} \times \exp\left(-\frac{\Delta\Delta_{\text{solv}}G^{\ddagger}}{RT}\right) \quad (2)$$

Equation (2) can be directly used to correct gas-phase bimolecular rate constants. The property $\Delta\Delta_{\text{solv}}G^{\ddagger}$ is the solvation Gibbs energy of activation, i.e., the difference between the solvation chemical potentials of the transition-state (TS) structure (denoted $\Delta_{\text{solv}}\bar{g}_{\text{TS,liq}}$) and those of the reactants ($\Delta_{\text{solv}}\bar{g}_{i,liq}$), as shown in equation (3) [13].

$$\Delta\Delta_{\text{solv}}G^{\ddagger} = \Delta_{\text{solv}}\bar{g}_{\text{TS,liq}} - \sum_i \Delta_{\text{solv}}\bar{g}_{\text{reactant } i,\text{liq}} \quad (3)$$

The solvation chemical potential of component i, $\Delta_{\text{solv}}\bar{g}_i$, is defined as the isothermal and isobaric change in the chemical potential when a molecule i, devoid of kinetic energy, is transferred from a perfect gas phase into a liquid phase [15,16].

To clarify, the relations between activation properties (denoted by symbol ‡) and solvation energy are graphically illustrated in Figure 1.



$$k_{gas}(T) = \frac{k_B T}{h}\exp\left(\frac{-\Delta G^{\ddagger}_{gas}}{RT}\right)$$

$$k_{liq} = k_{gas}\exp\left(\frac{-\Delta\Delta_{solv}G^{\ddagger}}{RT}\right)$$

Thus:

$$k_{liq} = \frac{k_B T}{h}\exp\left[\frac{-(\Delta G^{\ddagger} + \Delta\Delta_{solv}G^{\ddagger})}{RT}\right]$$

activation Gibbs energy in the liquid

Figure 1. *Potential energy surfaces associated with gas-phase and liquid-phase reactions. Graphical illustration of activation and solvation energies.*

There are many methods available to predict $\Delta_{\text{solv}}\bar{g}_i$, such as linear solvation energy relationships (LSERs) [17], machine-learning [17–19], continuum and discrete solvation models [20–22], and equations of state (EoS) [22,23].

The EoS approach, in particular, has already been successfully implemented in autoxidation kinetic simulations of pure n-alkanes using the UMR-PRU cubic EoS [24] to generate values of $\Delta_{\text{solv}}\bar{g}_i$ [11,25]. Nonetheless, in such simulations, the solvation Gibbs

energy of free radical species present in the kinetic mechanism was approximated as the solvation Gibbs energy of their parent molecules, i.e., their closed-shell version with a *H*-atom added on the radical site. As an example, ethyl hydroperoxide C_2H_5OOH is the parent molecule of the peroxyl radical C_2H_5OOg. However, this approximation turns out to be valid in non-polar liquid phase only (see Figure 3 of [11]), which can be an important source of error in simulations in polar media.

This approximation between free radicals and parent molecules seemed necessary because the pure compound and mixture inputs that are required by the cubic EoS (i.e., critical temperatures, critical pressures, acentric factors, and binary interaction parameters) are not available for free radicals. In addition, the existing models enabling to predict these input properties are limited to closed-shell molecules and, therefore, cannot be applied to radicals.

To overcome these limitations, the present work introduces new predictive methods to generate those inputs for free radicals based on simple and easy-to-use group contribution methods for the calculation of critical constants and acentric factors. Moreover, following a previous study of our group on the prediction of solvation energies of closed-shell molecules [26], a quantum-based continuum solvation model was incorporated into the *tc*-PR cubic EoS making it possible to predict the complex molecular interactions within mixtures (this solvation model relies on the well-established COSMO-RS approach).

Although the model proposed in this paper is totally new, some rare studies in the literature rely on a similar approach [27–30]. These studies have demonstrated that employing such a combination can yield reliable predictions of phase equilibria and solubility data across the entire fluid region of a phase diagram (liquid, vapor, and supercritical). A major advantage of this type of model is that it does not require binary interaction parameters as input and can therefore predict the thermodynamic properties of systems involving unstable or unusual molecules, such as free radicals, for which there is very little experimental data in the literature.

2 Modeling section

2.1.1 Predicting solvation quantities from an EoS

For a given temperature (T), pressure (P) and composition (molar fractions $\boldsymbol{z}$), the solvation chemical potential of a solute i in a liquid phase ($\Delta_{solv}\bar{g}_{i,liq}$) can be calculated with an equation of state (EoS) as follows [16,22,23]:

$$\Delta_{solv}\bar{g}_i\,(T,P,\mathbf{z}) = RT\ln\left[\frac{P\cdot V_m(T,P,\mathbf{z})\cdot\varphi_i\,(T,P,\mathbf{z})}{RT}\right] \tag{4}$$

This approach relies on the calculation of the fugacity coefficient of solute i (φ_i), and the estimation of the molar volume of the liquid phase (V_m). Both can be straightforwardly estimated by a cubic equation of state. To do this, we need first to calculate the EoS parameters of each pure compound (i.e., the attractive parameter, covolume and translation parameter, see their definitions below). Knowing these parameters, the mixture properties of interest can then be estimated at the specified temperature, pressure and composition. The following sections show how to calculate the properties of closed-shell molecules. A method adapted for free radicals is then proposed.

2.1.2 The Peng-Robinson EoS for pure molecular compounds (closed-shell molecules)

Cubic EoS are well-known for their good trade-off between accuracy and computational cost. In a previous study that dealt with solvation energies of closed-shell solutes [26], a state-

of-the-art version of the Peng-Robinson cubic EoS [31,32] was considered: the *tc*-PR EoS (*translated-consistent* Peng-Robinson) [33,34]. For a pure compound i, this EoS is defined as follows:

$$P = \frac{RT}{(V_m + c_i - b_i)} - \frac{a_i(T)}{(V_m + c_i)(V_m + c_i + b_i) + b_i(V_m + c_i - b_i)} \tag{5}$$

Where, for a compound i, $a_i(T)$ is the attractive parameter, b_i is the co-volume of the non-translated cubic EoS (simply called "co-volume" thereafter), and c_i is the volume translation. Based on the knowledge of the critical temperatures ($T_{c,i}$) and critical pressures ($P_{c,i}$), the co-volume (b_i) and the attractive parameter at the critical temperature ($a_i(T_{c,i})$ or simply $a_{c,i}$) can be obtained from equations (6) and (7).

$$b_i = \Omega_b \frac{RT_{c,i}}{P_{c,i}} \tag{6}$$

$$a_{c,i}(T) = \Omega_a \frac{\left(RT_{c,i}\right)^2}{P_{c,i}} \tag{7}$$

Where, Ω_a and Ω_b are universal constants depending on the version of the cubic EoS considered. For the Peng-Robinson EoS:

$$\begin{cases} \Omega_a = \dfrac{8(5X+1)}{49-37X} \approx 0.45724 \\ \Omega_b = \dfrac{X}{X+3} \approx 0.07780 \\ \text{with: } X = \left[1 + \sqrt[3]{4-2\sqrt{2}} + \sqrt[3]{4+2\sqrt{2}}\right]^{-1} \approx 0.25308 \end{cases} \tag{8}$$

The calculation of the parameter a_i for other temperatures than the critical one is done through the so called α-functions.

$$\begin{aligned} &a_i(T) = a_{c,i}\alpha_i(T_{r,i}) \\ &\text{with } T_{r,i} = T / T_{c,i} \text{ (reduced temperature)} \end{aligned} \tag{9}$$

α-functions are the core of every cubic EoS since they ensure its accuracy for a wide temperature range. Several α-functions are available in the literature, for instance, Soave [35], Boston-Mathias , Mathias-Copeman [36], and Twu 91 [37]. In the case of the *tc*-PR EoS, the Twu 91 α-function [37,38] was chosen among the other options due to its thermodynamic consistency (see details on this topic in [39]) and accuracy achieved by using compound-specific parameters fitted to experimental data (L_i, M_i, and N_i) [34]. The formulation of the Twu 91 α-function is given below.

$$\alpha_i(T_{r,i}) = T_{r,i}^{N_i(M_i-1)} \exp\left[L_i\left(1 - T_{r,i}^{M_i N_i}\right)\right] \tag{10}$$

When the parameters L_i, M_i, and N_i are not available, generalized correlations based on acentric factors (ω_i) can also be used instead [39]. Note that these correlations are not valid for quantum fluids such as H_2 and He.

$$L_i(\omega_i) = 0.0544 + 0.7536\omega_i + 0.0297\omega_i^2$$
$$M_i(\omega_i) = 0.8678 - 0.1785\omega_i + 0.1401\omega_i^2 \quad (11)$$
$$N_i = 2$$

A temperature-independent volume translation parameter (c_i) is added to improve the accuracy in predicting volumetric properties. The value of c_i is calculated to ensure that the equation of state replicates the experimental saturated liquid molar volume at a given temperature - in the present case, at $T_{r,i} = 0.8$. Optimal values for c_i are given in [34], but they can also be predicted by correlations based on critical constants and acentric factors, as shown in equation (12) [34,40].

$$c_i = \frac{RT_{c,i}}{P_{c,i}} \times \left(-0.014471 + 0.067498\omega_i - 0.084852\omega_i^2 + 0.067298\omega_i^3 - 0.017366\omega_i^4\right) \quad (12)$$

2.1.3 Calculation of mixture EoS parameters from advanced mixing rules

For dealing with mixtures, the formulation of the cubic EoS is essentially the same, see equation (5), but the parameters $a_i(T)$, b_i, and c_i become mixture parameters ($a(T,\boldsymbol{z})$, $b(\boldsymbol{z})$, and $c(\boldsymbol{z})$) that are functions of the composition of the system ($\boldsymbol{z}$). The pure compound parameters and the mixture parameters are related through the so-called mixing rules. Inspired from the works of Huron and Vidal, and Michelsen [41,42], the following advanced mixing rules were proposed for NC-component mixtures [26,43,44].

$$b(\mathbf{z}) = \sum_{i=1}^{NC}\sum_{j=1}^{NC} z_i z_j b_{ij}$$
$$\text{with } \begin{cases} b_{ij} = \left(\dfrac{b_i^{1/s} + b_j^{1/s}}{2}\right)^s \\ s = 3/2 \end{cases} \quad (13)$$

$$c(\mathbf{z}) = \sum_{i=1}^{NC} z_i c_i \quad (14)$$

$$\frac{a(T,\mathbf{z})}{b(\mathbf{z})} = \frac{g_{res}^{E,\gamma}(T,\mathbf{z})}{\Lambda_{EoS}} + \sum_{i=1}^{NC} z_i \frac{a_i(T)}{b_i} \quad (15)$$
$$\text{with } \Lambda_{EoS} = -0.6232$$

The mixture parameters $b(\boldsymbol{z})$ and $c(\boldsymbol{z})$ depend only on the pure compound parameters (b_i and c_i), and the mixture composition $\boldsymbol{z}$. The calculation of the attractive parameter of the mixture (a/b) requires the estimation of the residual Gibbs excess energy ($g_{res}^{E,\gamma}$), which is related to the residual activity coefficient of each compound in the mixture (γ_i^{res}), as shown in equation (16) [43]. It is recalled here that the *residual* part of an activity coefficient model refers to its enthalpic (or non-combinatorial) contribution. Any activity coefficient model with a residual contribution can be used in this equation. As in a previous study [26], COSMO-RS was used to calculate residual activity coefficients.

$$\frac{g_{res}^{E,\gamma}(T,\mathbf{z})}{RT} = \sum_{i=1}^{NC} x_i \ln \gamma_i^{res}(T,\mathbf{z}) \tag{16}$$

In the COSMO-RS framework, the prediction of residual activity coefficients depends exclusively on pure compound information, which stands for the probability distribution of the polarization charge density (σ) over molecular cavity surfaces [45]. This distribution, obtained through quantum calculations, is represented in a histogram called σ-profile (denoted as $p_i(\sigma)$ for a pure compound i). An example of σ-profile is given in Figure 2.

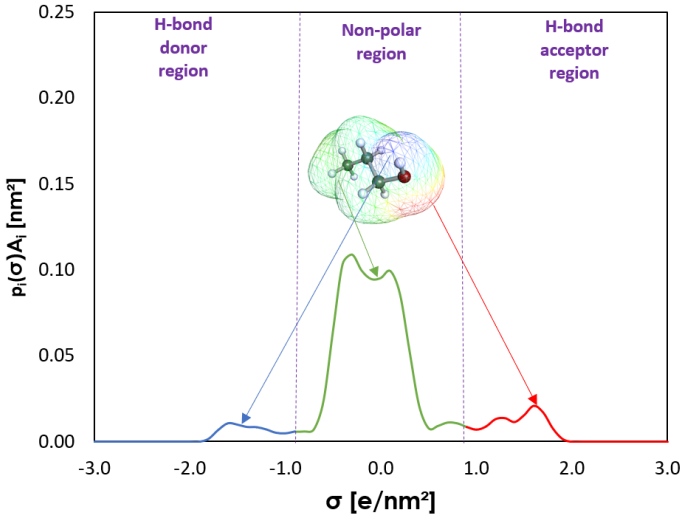


Figure 2. *Example of a σ-profile histogram (1-propanol), where A_i is the total COSMO cavity area, $p_i(\sigma)$ the probability distribution of the polarization charge density (σ).*

As it requires neither experimental data nor fitted binary interaction parameters [46], the COSMO-RS model is particularly interesting for predicting activity coefficients involving free radicals. Details of the formulation of the COSMO-RS version of the model used in the mixing rule are given elsewhere [26].

2.2 Adapting the EoS approach for dealing with free radicals

All the procedure presented above can be applied to closed-shell molecules only, since the pure compound parameters (a_i, b_i, and c_i) cannot be estimated for free radicals due to a lack of knowledge of critical properties ($T_{c,i}$ and $P_{c,i}$), and acentric factors (ω_i) for such species. For this reason, a new method was developed in this work to enable the prediction of the pure compound parameters without the *a priori* knowledge of $T_{c,i}$, $P_{c,i}$ and ω_i. In addition, a fast method to predict σ-profiles of these radicals is also proposed.

2.2.1 Proposal of group contribution methods to predict T_c, P_c, and ω of molecules and free radicals

The starting point for the following discussion is the empirical observation that power functions of both the co-volume (b_i) and the attractive parameter at the critical point ($a_{c,i}$) show a linear dependence on molecular weight (MW_i) for a homologous series of compounds, as pinpointed in Figure 3. This linear behavior tends to prove that b_i and $a_{c,i}$ can be predicted by the concept of group additivity. Based on this, a simple and easy-to-use group contribution method is proposed in equations (17) and (18).

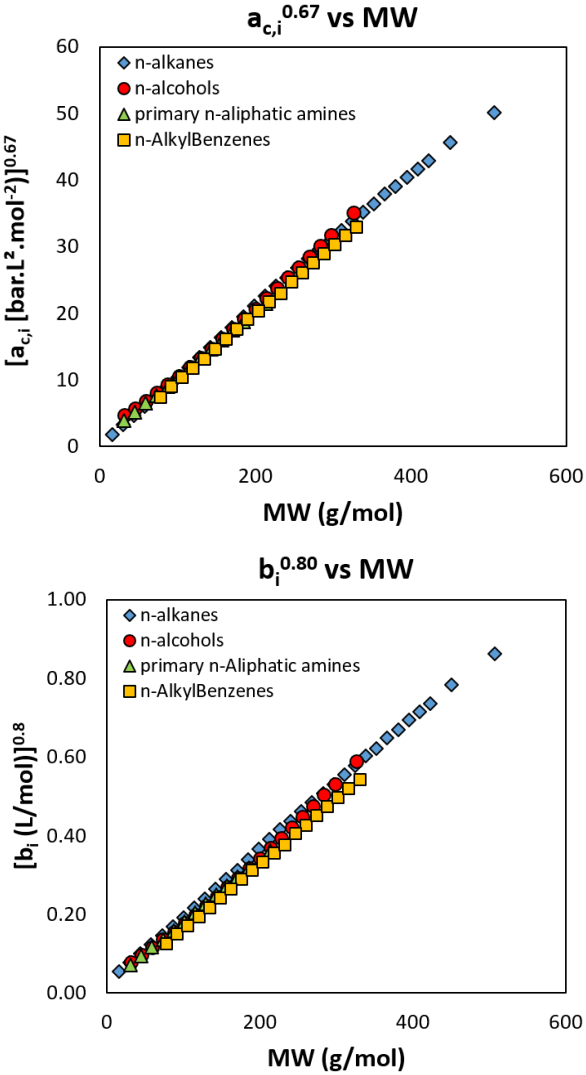


Figure 3. *Linear dependence of power functions of $a_{c,i}$ and b_i with respect to the molecular weight (MW) of pure compounds sorted by homologous family. Note that the linearity is achieved when the exponents of b_i* and *$a_{c,i}$ are 0.80 and 0.67, respectively.*

$$\left(\frac{a_{c,i}}{\Omega_a R^2}\right)^{\beta_{ac}} = \left(\frac{T_{c,i}^2}{P_{c,i}}\right)^{\beta_{ac}} = \sum_{j=1}^{n_g} N_{g_j} \cdot A_{c,[g_j]} \tag{17}$$

$$\left(\frac{b_i}{\Omega_b R}\right)^{\beta_b} = \left(\frac{T_{c,i}}{P_{c,i}}\right)^{\beta_b} = \sum_{j=1}^{n_g} N_{g_j} \cdot B_{[g_j]} \tag{18}$$

In the equations above, N_{g_j} is the occurrence of the group g_j in the molecule i, whereas $A_{c,[g_j]}$ and $B_{[g_j]}$ are the contributions of the functional group g_j to the properties $a_{c,i}$ and b_i, respectively. The parameters β_{ac} and β_b are the exponents of the power functions of $a_{c,i}$ and b_i, and they turn to be 0.67 and 0.80, respectively (more details on this will be given in Appendix 1). Note that the expressions of $a_{c,i}/(\Omega_a R^2)$ and $b_i/(\Omega_b R)$ with respect to $T_{c,i}$ and $P_{c,i}$ come from equations (6) and (7). The methodology used for the determination of the

group contributions $A_{c,[g_j]}$ and $B_{[g_j]}$ for molecular species and free radicals is detailed in section 3.2.

Once $a_{c,i}$ and b_i are known, the critical temperature ($T_{c,i}$) and pressure ($P_{c,i}$) of a molecule or a free radical can be estimated by equations (19) and (20) which were derived from equations (6) and (7).

$$T_{c,i} = \left(\frac{\Omega_b}{\Omega_a}\frac{1}{R}\right)\frac{a_{c,i}}{b_i} \tag{19}$$

$$P_{c,i} = \left(\frac{\Omega_b^2}{\Omega_a}\right)\frac{a_{c,i}}{b_i^2} \tag{20}$$

To summarize, at this step of the procedure, the proposed group-contribution method allows the determination of the critical attractive parameter $a_{c,i} = a(T_{c,i})$, the co-volume b_i, as well as the critical properties $T_{c,i}$ and $P_{c,i}$ of radicals.

It now remains to determine the attractive parameter of free radicals for temperatures other than the critical one. As mentioned above, this is done through the use of an α-function, as shown in equation (9). However, determining the α-function value at any temperature without having, at least, a good estimation of the acentric factor (ω_i) is a tricky task. When experimental vapor-pressure data are available, the best and most robust option is to fit the parameters involved in the α-function to these data. Since vapor-pressure data of free radicals cannot be measured, it is proposed to consider the Soave α-function which shows the simple formulation given by equation (21).

$$\alpha(T_{r,i}) = \left[1 + m_i\left(1 - \sqrt{T_{r,i}}\right)\right]^2 \quad \text{with: } T_{r,i} = T / T_{c,i} \tag{21}$$

Such an α-function contains a single component-dependent parameter (m_i), estimated – as a general rule – from a generalized correlation based on the acentric factor (ω_i) [47]. As mentioned above, the acentric factors of free radicals are never known and therefore, another approach has to be proposed. In this study, the idea is to estimate m_i by group contribution. In order to achieve this goal, a specific group contribution method has been developed with the aim of estimating not m_i but the value of the attractive parameter a_i at 0 K (see equation (22)). The reason for this is that $a_i(0\,\text{K})$ and m_i are directly related and that $a_i(0\,\text{K})$ is more suitable than m_i for the application of the group contribution concept.

$$\left(\frac{a_i(0)}{\Omega_a R^2}\right)^{\beta_{a0}} = \sum_{j=1}^{n_g} N_{g_j} \cdot A_{0,[g_j]} \tag{22}$$

Where, $A_{0,[g_j]}$ is the contribution of the functional group g_j for the calculation of $a_i(0)$. Note that a group contribution method can be envisaged because a linear behavior is obtained when a power function of $a_i(0)$ is plotted as a function of the molecular weight for most of homologous series of compounds, as illustrated in Figure 4. Note also that in this particular instance, the exponent of the power function, β_{a0}, is equal to 0.48, and the rationale behind this choice is elaborated in Appendix 1.

As mentioned above, the methodology used for the determination of the group contributions $A_{0,g}$ is detailed in section 3.2.

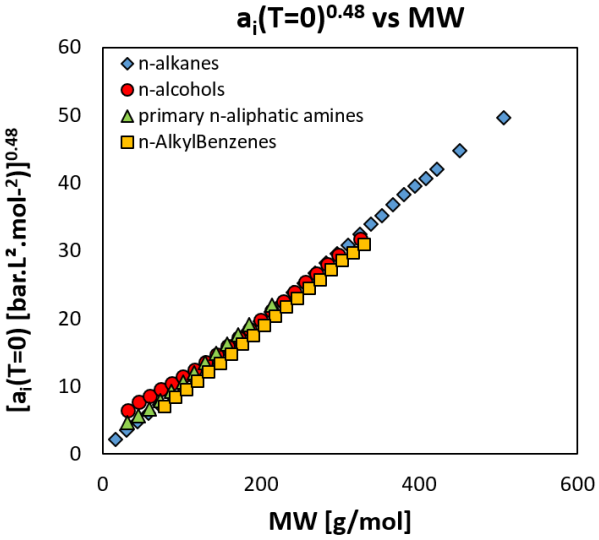


Figure 4. *Linear dependence of power functions of $a_i(0)$ with respect to the molecular weight (MW) of pure compounds sorted by homologous family. Note that the linearity is achieved for an exponent value of $a_i(0)$ equal to 0.48.*

By combining Eqs. (9) and (21) and by setting the reduced temperature T_r to 0, the expression for the parameter m_i is obtained.

$$m_i = \sqrt{\frac{a_i(0)}{a_{c,i}}} - 1 \tag{23}$$

Where both $a_i(0)$ and $a_{c,i}$ are calculated by group contribution (equations (17) and (22), respectively). Once m_i is known, the α-function can be calculated for any temperature by means of equation (21). Note that the critical temperature ($T_{c,i}$), required for such a calculation, is obtained from equation (19). Once $a_{c,i}$, b_i and m_i are known, one can perform pure compound calculations using the PR-EoS based on the Soave alpha-function.

At this step, we recall that the acentric factors are input parameters used in generalized correlations for the volume translation parameter, c_i [40] (see, e.g., Eq. (12)), which improves the molar volume prediction of liquid solvents. For the majority of solvents, acentric factors are readily available from experimental databases. When dealing with unconventional solvents, whose critical properties and acentric factors are not readily available, acentric factors must be predicted. The prediction of the acentric factor relies on its definition, as outlined in the equation (24).

$$\omega_i = -\log_{10}\left[\frac{P_i^{sat}(T = 0.7T_{c,i})}{P_{c,i}}\right] - 1 \tag{24}$$

Where, P_i^{sat}, the vapor pressure of pure component i, can be calculated with the non-translated version of the PR-EoS at a reduced temperature of 0.7 (i.e., $T/T_{c,i} = 0.7$). The $T_{c,i}$ and $P_{c,i}$ properties can be obtained by group contribution using equations (19) and (20).

Note that for highly diluted solutes (mole fraction close to zero), their volume translation (c_i) has no impact on the calculation of the total volume translation of the liquid phase (see equation (14)). For closed-shell solutes, if the acentric factor is known, the volume correction can be calculated by the correlation presented in equation (12). For the sake of simplicity, the

volume translation of a free radical ($c_{X\bullet}$) can be taken to be equal to that of its parent molecule (c_{XH}):

$$c_{X\bullet} = c_{XH} \tag{25}$$

For parent molecules (and consequently the free radicals associated) whose acentric factor is unknown, the volumetric correction can be assumed to be zero, without impacting significantly the accuracy of the method (this correction is essentially relevant for solvent molecules). In this study, all the acentric factors of the parent molecules were known, so it was possible to calculate the volumetric correction, c_{XH}, for all of them.

2.2.2 Fast prediction of σ-profiles of molecules and free radicals

σ-profiles are the inputs used for establishing the connection between the attractive parameters of pure components and the attractive parameter of the corresponding mixture through the calculation of residual activity coefficients that appear in the mixing rule (see equation (15)). As previously stated, σ-profiles are originally obtained by quantum chemistry calculations, which are often very time-consuming. This is why the σ-profiles of usual molecules are generally stored in databases. For instance, the database provided with the software *COSMOTherm* (version 2016 onwards) includes more than 10,000 molecules.

For the molecules that are not present in this database, the quantum-chemical calculations need to be performed. An option to shortcut the need for quantum-based calculation is to use a group contribution approach [26,48,49], which is less accurate, but provides a much faster generation of σ-profiles.

$$p_i(\sigma_m) = \sum_{j=1}^{n_g} N_{g_j} \cdot p_{[\mathrm{g_j}]}(\sigma_m)$$

$$\text{For discrete values of } \sigma_m \in [-3.0, +3.0] \text{ e/nm}^2 \tag{26}$$

$$\text{Step: } \Delta\sigma = 0.1 \text{ e/nm}^2$$

In this study, the GC-based approach to predict σ-profiles was extended to free radicals by simply adding functional groups containing radical sites.

2.2.3 Decomposition scheme into functional groups

For the sake of simplicity, the proposed GC-methods rely on a first-order group decomposition scheme (second or third order methods take into account the chemical environment of the 1st order molecular groups and are much more complex to be implemented).

Basically, four different group contribution methods are proposed in this study, for the calculation of $a_{c,i}$, b_i, $a_i(0)$, and $p_i(\sigma_m)$. All these methods share the same group decomposition scheme based on UNIFAC functional groups. These groups were divided into two categories:

1. Conventional UNIFAC groups (g). Example: $[g] = [CH_3]$.
2. Specific UNIFAC groups $[g^\bullet\ (-)\ g]$ to mimic the loss of a hydrogen atom by a given molecule ($[g^\bullet\ (-)\ g]$ must be read "$g^\bullet$ minus g" and therefore, is equivalent to a *negative group* "minus H"). Such a group needs to be added to the decomposition in elementary groups of the parent molecule to describe the corresponding free radical. It simply can be seen as a *correction* to bring to the parent molecule. An example of such a group is: $[g^\bullet\ (-)\ g] = [CH_2^\bullet\ (-)\ CH_3]$.

Therefore, in this framework, the decomposition of a free radical $X^{\bullet}$ into functional groups is equal to the decomposition of its parent molecule *XH* in addition to the correction group for free radicals. As an example, let us consider the free radical formed from a generic molecule with a known decomposition. The radical formation reaction is given in Figure 5.

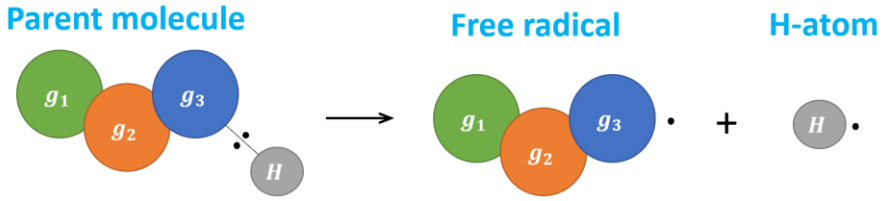


Figure 5. *Generic reaction for the formation of a free radical.*

The decomposition of both parent molecule and free radical would be as follows:

Parent molecule (XH): $[g_1]$:1, $[g_2]$:1, $[g_3]$:1

Free radical ($X^{\bullet}$): $[g_1]$:1, $[g_2]$:1, $[g_3]$:1, $[g_3^g(-)g_3]$:1

2.2.4 *Recipe for applying the proposed group contribution methods to free radicals*

To perform mixture calculations with the proposed equation of state, we need to provide the following pure compound inputs for every compound in the system:

- Attractive parameter at T: $a_i(T) = a_{c,i} \cdot \alpha(T, m_i)$
- Co-volume: b_i
- Volume translation: c_i
- σ-profile: $p_i(\sigma)$

For the generic free radical of Figure 5, these parameters are calculated using the group contribution methods introduced above. First of all, the attractive parameter at the critical point ($a_{c,i}$), the co-volume (b_i) and the attractive parameter at $T = 0$ ($a_i(0)$), of such a radical are calculated following equations (17), (18) and (22), as follows:

$$\underbrace{\left(\frac{a_{c,X^{\bullet}}}{\Omega_a R^2}\right)^{\beta_{ac}}}_{\text{free radical}} = \underbrace{1\times A_{c,[g_1]} + 1\times A_{c,[g_2]} + 1\times A_{c,[g_3]}}_{\text{parent molecule}} + \underbrace{1\times A_{c,[g_3^g(-)g_3]}}_{\text{correction}} \tag{27}$$

$$\underbrace{\left(\frac{b_{X^{\bullet}}}{\Omega_b R}\right)^{\beta_b}}_{\text{free radical}} = \underbrace{1\times B_{[g_1]} + 1\times B_{[g_2]} + 1\times B_{[g_3]}}_{\text{parent molecule}} + \underbrace{1\times B_{[g_3^g(-)g_3]}}_{\text{correction}} \tag{28}$$

$$\underbrace{\left(\frac{a_{X^{\bullet}}(0)}{\Omega_a R^2}\right)^{\beta_{a0}}}_{\text{free radical}} = \underbrace{1\times A_{0,[g_1]} + 1\times A_{0,[g_2]} + 1\times A_{0,[g_3]}}_{\text{parent molecule}} + \underbrace{1\times A_{0,[g_3^g(-)g_3]}}_{\text{correction}} \tag{29}$$

What is interesting about the proposed decomposition scheme is that the part of the group contribution method that corresponds to the parent molecule can be calculated using experimental inputs of $T_{c,XH}$, $P_{c,XH}$, and ω_{XH}, as shown below. By doing so, the accuracy of the method is expected to be increased, especially for small molecules.

$$\underbrace{\left(\frac{a_{c,X\bullet}}{\Omega_a R^2}\right)^{\beta_{ac}}}_{\text{free radical}} = \underbrace{\left(\frac{T_{c,XH}^2}{P_{c,XH}}\right)^{\beta_{ac}}}_{\text{parent molecule}} + \underbrace{1\times A_{c,[g_3^g(-)g_3]}}_{\text{correction}} \tag{30}$$

$$\underbrace{\left(\frac{b_{X\bullet}}{\Omega_b R}\right)^{\beta_b}}_{\text{free radical}} = \underbrace{\left(\frac{T_{c,XH}}{P_{c,XH}}\right)^{\beta_b}}_{\text{parent molecule}} + \underbrace{1\times B_{[g_3^g(-)g_3]}}_{\text{correction}} \tag{31}$$

$$\begin{cases} \underbrace{\left(\frac{a_{X\bullet}(0)}{\Omega_a R^2}\right)^{\beta_{a0}}}_{\text{free radical}} = \underbrace{\left(\frac{T_{c,XH}^2\cdot\alpha_{XH}(0)}{P_{c,XH}}\right)^{\beta_{a0}}}_{\text{parent molecule}} + \underbrace{1\times A_{0,[g_3^g(-)g_3]}}_{\text{correction}} \\ \text{with: } \alpha_{XH}(0) = (1+m_{XH})^2 \text{ and } m_{XH} = 0.37464 + 1.54226\omega_{XH} - 0.26992\omega_{XH}^2 \end{cases} \tag{32}$$

Note that the parent molecule must be described by the Soave α function, which predicts a finite limit for $a(0)$, and not by the Twu 91 function which diverges at 0 K.

The critical temperature of the free radical, $T_{c,X\bullet}$, can be obtained from $a_{c,X\bullet}$ and $b_{c,X\bullet}$, using equation (19), the $m_{X\bullet}$ coefficient can be deduced from equation (23), which, in turn, enables the calculation of the free radical α-function at a given temperature T by means of equation (21). It is recalled that, for simplicity, the volume translation parameter of the free radical ($c_{X\bullet}$), is approximated as being equal to that of its parent molecule (c_{XH}), which can be obtained, for instance, by the generalized correlation given in equation (12).

Finally, a GC method for predicting the σ-profile of the free radical, $p_{X\bullet}(\sigma_m)$, based on a correction to be made to the parent molecule $p_{XH}(\sigma_m)$, is applied, as shown in equation (33).

$$\underbrace{\left[p_{X\bullet}(\sigma_m)\right]}_{\text{free radical}} = \underbrace{\left[p_{XH}(\sigma_m)\right]}_{\text{parent molecule}} + \underbrace{\left[\Delta p_{[g_3^g(-)g_3]}(\sigma_m)\right]}_{\text{correction}} \tag{33}$$

In the equation above, the parent molecule σ-profile is obtained from a database of quantum-calculated σ-profiles and the method will be referred to in the remainder of the text as a *GC-based correction* method.

If a parent molecule is not present in the database, the group contribution method developed in [26] can be used instead, as shown in equation (34). A slight degradation of the results could be observed in such a case.

$$\underbrace{\left[p_{X\bullet}(\sigma_m)\right]}_{\text{free radical}} = \underbrace{\left[1\times p_{[g_1]}(\sigma_m) + 1\times p_{[g_2]}(\sigma_m) + 1\times p_{[g_3]}(\sigma_m)\right]}_{\text{parent molecule}} + \underbrace{\left[\Delta p_{[g_3^g(-)g_3]}(\sigma_m)\right]}_{\text{correction}}$$

For discrete values of $\sigma_m \in [-3.0, +3.0]$ e/nm² (34)

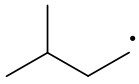
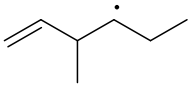
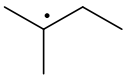
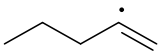
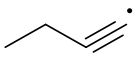
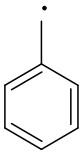
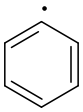
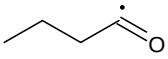
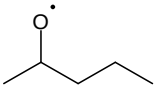
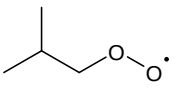
Step: $\Delta\sigma$ = 0.1 e/nm²

3 Technical details

3.1 Functional groups considered in this work

The set of g-type groups is the same as that used in a previous publication [26], which comprises 134 UNIFAC groups. For the specific groups involving radical sites $[g^{\bullet}\ (-)\ g]$, ten types of free radicals involved in oxidation mechanisms were considered, as summarized in Table 1.

Table 1. *Types of C/H/O free radicals considered in this study.*

#	Radical category	Example of molecule	Specific groups (correction) for radicals
1	Primary		$[CH_2^{\bullet}\ (-)\ CH_3]$
2	Secondary		$[CH^{\bullet}\ (-)\ CH_2]$
3	Tertiary		$[C^{\bullet}\ (-)\ CH]$
4	Vinyl		$[CH_2=C^{\bullet}\ (-)\ CH_2=CH]$
5	Acetylenic		$[C\equiv C^{\bullet}\ (-)\ C\equiv CH]$
6	Benzyl		$[AC-CH_2^{\bullet}\ (-)\ AC-CH_3]$
7	Phenyl		$[AC^{\bullet}\ (-)\ ACH]$
8	Carbonyl		$[R-C^{\bullet}=O\ (-)\ R-CH=C$
9	Alkoxyl		$[R-O^{\bullet}\ (-)\ R-OH]$
10	Peroxyl		$[R-O-O^{\bullet}\ (-)\ R-O-OH$

Based on the C/H/O radical types presented in Table 1, 42 specific groups $[g^{\bullet}\ (-)\ g]$ were identified. The reader is referred to Appendix 2 to see the groups (g) and corrections $[g^{\bullet}\ (-)\ g]$, available. The parameter values for each group are given in the supplementary material.

3.2 Parametrization of the group contribution methods

3.2.1 σ-profiles

Based on the group decomposition of closed-shell molecules and free radicals, a multilinear regression algorithm was implemented to determine the contribution of each group g and *correction* $[g^{\bullet}(-)g]$ in the calculation of σ-profiles. The quantum-based σ-profiles of 1613 molecules (from the COSMOTherm database v.2016) and 135 free radicals (generated in this work with *Turbomole* v.2016) were used as reference for such a regression. These σ-profiles were obtained at the BP-TZVPD-FINE level of theory. It should be noted that only the most stable conformer configuration of the parent molecule was considered.

3.2.2 EoS parameters and critical constants

For closed-shell molecules, the optimization of the group contribution parameters, i.e., $A_{c,[g_i]}$, $B_{[g_i]}$, and $A_{0,[g_i]}$, were determined through a multilinear regression algorithm based on the experimental values of $T_{c,i}$, $P_{c,i}$, and ω_i of the set of 1613 molecules used for the prediction of σ-profiles.

For radical species, the same method could not be implemented due to the lack of experimental critical data and acentric factors. Consequently, group parameters were fitted differently:

1. Following the work of Hsieh and Lin [50] in which the parameter b_i of molecules was correlated to the volume of their cavity (V_i^{cav}) within a continuum dielectric medium (multiplied by Avogadro's number, N_A) [51], a multilinear regression algorithm was used to calculate the contributions of the 134 UNIFAC groups ($V_{[g_i]}$) and the 42 corrections for radicals ($V_{[g^g(-)g]}$), using the same power that was applied to the b_i parameter (equation (18)), but in this case using cavity volumes V_i^{cav}, as shown in the equation (35). For this regression, the cavity volume of 1613 molecules and 135 radicals were considered. Therefore, the free radical correction was calculated by simply making $B_{[g^g(-)g]} = V_{[g^g(-)g]}$. Note that, for the UNIFAC groups, the $B_{[g_i]}$ values fitted from b_i data obtained experimental values of $T_{c,i}$, $P_{c,i}$, and $\omega_i$were retained.

$$\left(\frac{N_A V_i^{cav}}{\Omega_b R}\right)^{\beta_b} = \sum_{j=1}^{n_g} N_{g_j} \cdot V_{[g_j]} \tag{35}$$

2. As for the corrections $A_{c,[g^{\bullet}(-)g]}$ and $A_{0,[g^{\bullet}(-)g]}$ (84 parameters from 42 groups), they were fitted to the solvation Gibbs energy data using a Nelder-Mead optimization algorithm. For this purpose, a pseudo-experimental database for the solvation of the 135 free radicals infinitely diluted in 78 pure solvents were generated using the software *COSMOTherm*, also at the BP-TZVPD-FINE level of theory. The temperature range considered was between 298 and 398 K, resulting in a databank containing 52260 data points.

The calculations with *COSMOTherm* used the full set of conformers for the solvents and only the most stable conformer for the solutes (free radicals). We checked that using only the most stable conformer for the solute molecule does not significantly reduce the accuracy of the model. On average, the deviations are less than 0.1 kcal/mol, as it will be seen in section 4.2.2.

For adjusting the parameters of free radical groups used in the cubic EoS, the objective function OF given in equation (36) was employed.

$$OF = \sum_{k=1}^{N} \left[\left(\Delta_{\text{solv}} \bar{g}_{X^\bullet}^{\infty} \right)_{\text{k}}^{\text{COSMOTherm}} - \left(\Delta_{\text{solv}} \bar{g}_{X^\bullet}^{\infty} \right)_{\text{k}}^{\text{EoS}} \right]^2 \tag{36}$$

Where $\left(\Delta_{\text{solv}} \bar{g}_{X^\bullet}^{\infty} \right)_{\text{k}}^{\text{COSMOTherm}}$ is the value of the solvation Gibbs energy of the radical $X^\bullet$ obtained with the *COSMOTherm* software and $\left(\Delta_{\text{solv}} \bar{g}_{X^\bullet}^{\infty} \right)_{\text{k}}^{\text{EoS}}$ is the solvation Gibbs energy of the same radical calculated with the proposed EoS, for the k-th of N experimental points.

Note that the reference data generated by *COSMOTherm* do not require pressure as an input since it uses a pressure independent continuum solvation model to perform the calculations (COSMO-RS). However, both temperature and pressure conditions must be set when using an equation of state approach to predict solvation properties. Knowing that:

- The solvation Gibbs energies of the components of a liquid mixture vary little with pressure due to the incompressible nature of liquids,
- At a given temperature, a mixture containing a solute infinitely diluted in a solvent is liquid at a pressure higher than the vapor pressure of the solvent.

Consequently, we assumed that each of the isothermal mixing data generated by *COSMOTherm* was at a pressure equal to the vapour pressure of the solvent. The calculations with the cubic EoS were carried out by an in-house *MATLAB* routine, with the COSMO-RS parametrization published in a previous work [26]. In the optimization process, the pure compound parameters of free radicals were calculated through equations (30) to (33). Note that experimental data were used for the critical temperature (T_c), critical pressure (P_c,), and acentric factor (ω) of parent molecules. Additionally, quantum-based σ-profiles of parent molecules were considered in the calculation of free radicals σ-profiles through equation (33). It was stipulated that for each radical correction group there were at least 3 free radicals in the database. Note that some groups correspond to entire molecules, and therefore, they have only 1 representative radical. For example, the correction for the free radical formed from methane, $[CH_3^{\bullet}(-)CH_4]$, is applicable only to this free radical, in particular.

3.3 Deviations analysis

To assess the deviations in predicting the cubic EoS parameters, the critical constants, and acentric factors for closed-shell molecules, we employed the Mean Percentage Error (MPE). This statistical measure is calculated by equation (37), in which y_i is the predicted value of the property, $\hat{y}_i$ is its reference value (experimental from the DIPPR database). N is the size of the dataset.

$$MPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| \tag{37}$$

Given that solvation Gibbs energy values often approach zero, employing absolute deviations becomes more apt for this quantity. In this instance, we relied on the mean unsigned error (MUE) for the deviation evaluation, whose definition is given in (38). In this case, y_i is the values of $\Delta_{\text{solv}} \bar{g}_i$ calculated by the equation of state and $\hat{y}_i$ is its reference value (experimental from the CompSol database for closed-shell molecules or calculated with *COSMOTherm* for free radicals).

$$MUE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \quad (38)$$

4 Results

4.1 GC-methods applied to closed-shell molecules

4.1.1 Prediction of critical constants and EoS parameters

Figure 6 (panels a and b) compares the PR EoS pure compound parameters, $a_{c,i}$ and b_i, obtained using experimental critical inputs (abscissa axis in the parity plots) and the GC-method proposed in this study (ordinate axis), for 1613 closed-shell molecules. The calculated results are in good agreement with the experimental ones, since the mean percentage errors for all data is below 8% for both EoS parameters.

It was found that the method provided satisfactory predictions of $a_{c,i}$ and b_i for most families of compounds. Larger deviations were obtained for organic salts, some cyclic and aromatic compounds, polyols, glycerides, as well as, some carboxylic acids and esters, especially the polyfunctional ones. The detailed deviation results by chemical families are given in the supplementary material.

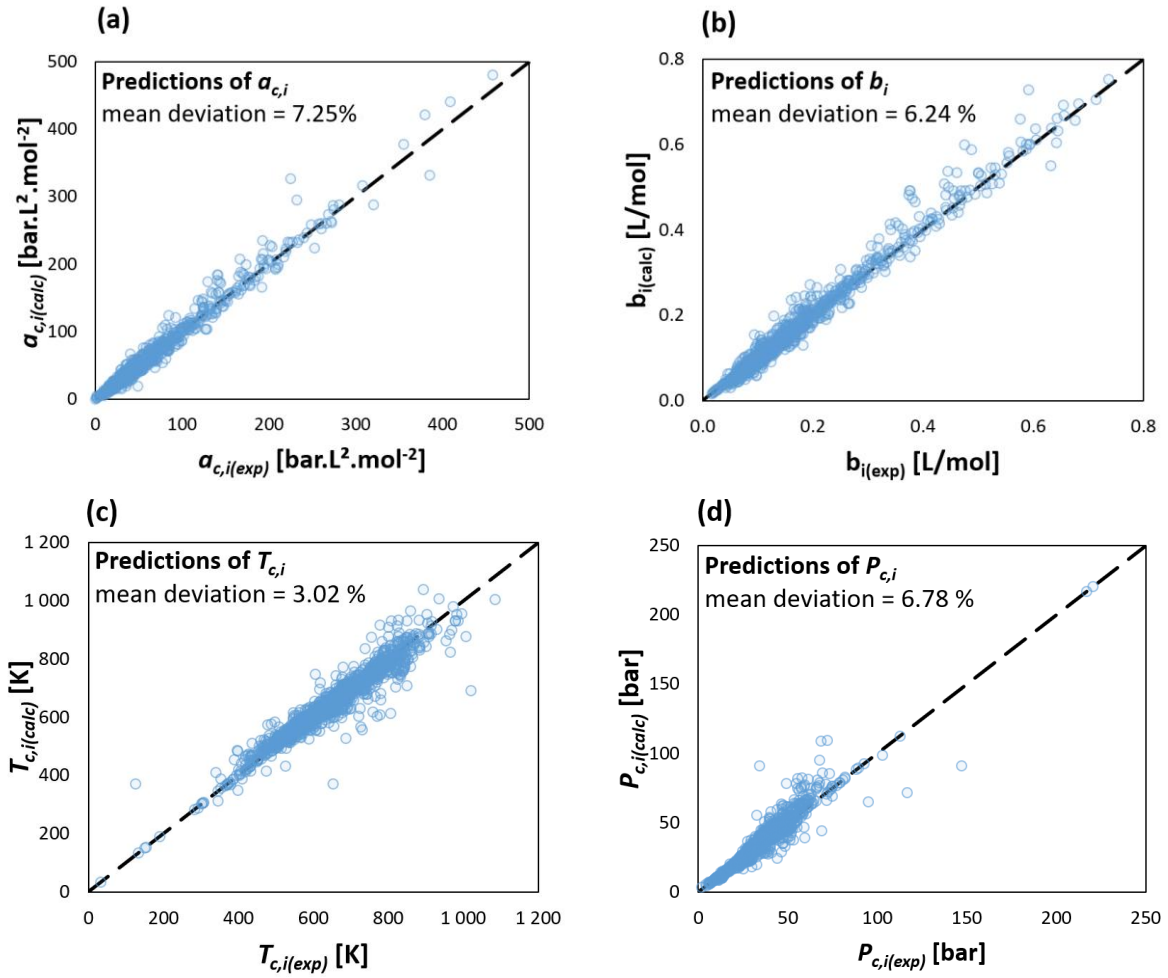


Figure 6. *Parity plots for the predictions of the PR-EoS parameters ($a_{c,i}$ and b_i) and critical constants ($T_{c,i}$ and $P_{c,i}$) of 1613 closed-shell molecules using the group contribution methods proposed in this work.*

Once the parameters $a_{c,i}$ and b_i have been determined, it becomes possible to derive the critical constants of the molecules ($T_{c,i}$ and $P_{c,i}$) using equations (19) and (20). The parity plots for the prediction of the critical constants are also given in Figure 6. Again, a satisfactory agreement between predicted and experimental data is obtained, especially for $T_{c,i}$. As shown in Table 2, the accuracy achieved with the proposed first-order GC method is lower than with more sophisticated predictive models based on second and third-order GC approaches or machine learning (ML), but on the other hand, the dataset considered here is the much larger one.

Table 2. *Mean Percentage Errors (MPE) for different models in the prediction of critical temperatures and pressures.*

Reference	Method	Dataset ($T_{c,i}/P_{c,i}$)	$MPE_{T_{c,i}}$(%)	$MPE_{P_{c,i}}$(%)
[52]	2nd order GC	285 / 289	0.85	2.89
[53]	3rd order GC	587 / 573	0.90	2.20
[54]	ML	917 / 684	0.40	5.00
[55]	ML	888 / 782	1.30	3.82
This work	1st order GC	1613 / 1613	3.02	6.78

If we now shift our focus to the prediction of the Soave α-function parameter (m_i) via the calculation of $a_i(0)$, we can see in Figure 7 that reasonably accurate predictions of this α-function parameter can be made, with deviations remaining below 8% on average. Additionally, Table 3 provides a comparative analysis of the acentric factor predictions using diverse models. Note that, with respect to our proposed approach, the acentric factors were calculated in accordance with its definition (see equation (24)).

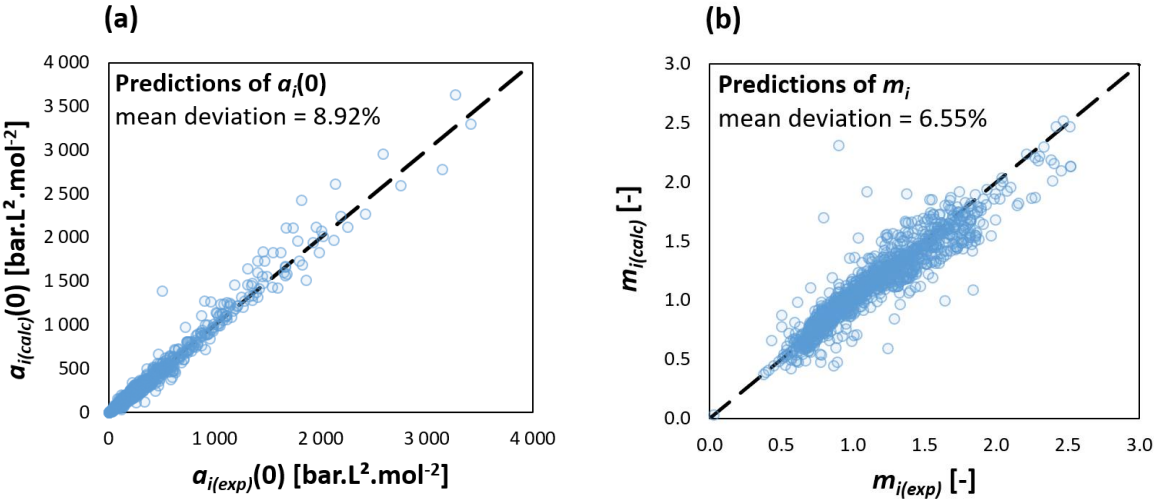


Figure 7. *Parity plots for the predictions of $a_i(0)$ and m_i of 1613 closed-shell molecules using the group contribution methods proposed in this work.*

Table 3. *Mean Percentage Errors (MPE) for different models in the prediction of acentric factors.*

Reference	Method	Dataset ($\omega_{c,i}$)	$MPE_{\omega_{c,i}}$(%)
[56]	2nd order GC	181	2.99
[54]	ML	614	6.90
[55]	ML	524	8.74
This work	1st order GC	1613	12.9

Predicting acentric factors poses a formidable challenge due to their reliance on the size and polarizability of molecules. Table 3 shows that the differences between experimental and

predicted acentric factors are significantly larger than those associated with predictions of critical constants such as $T_{c,i}$ and $P_{c,i}$. As expected, the first-order group contribution approach exhibited greater deviations in comparison to more advanced methods. This result must be balanced by the larger data set used here to fit the group parameters compared to other methods.

In any event, it is important to emphasize that the comparison results presented in Table 2 and Table 3 provide only a preliminary indication of the accuracy of each approach. A more comprehensive evaluation would require testing the methods on the same dataset.

4.1.2 Predictions of solvation Gibbs energy of closed-shell molecules

In order to validate the proposed group contribution methods for the prediction of solvation Gibbs energy data, we first considered only the solvation of closed-shell molecules, for which experimental data are available in the literature [16]. Four different versions of the proposed equation of state were evaluated, as shown in Table 4.

- The first two versions (V1 and V2) are semi-predictive. Both of them use experimental inputs to estimate the pure component EoS parameters ($a_{c,i}, b_i$, α function parameters, c_i). The difference between versions V1 and V2 lies in the way the σ-profiles of the molecules are estimated: V1 considers a quantum-chemical approach while V2 uses a GC method. It is recalled that the σ-profiles influence the COSMO-RS calculations and in particular the estimation of the residual activity coefficients (see Eq. (16)) which are needed to estimate the mixture EoS parameters.
- The last two versions (V3 and V4) use fully predictive methods for the estimation of pure-component EoS parameters. Similar to V1 and V2, the only difference between these versions concerns the way the σ-profiles are estimated (quantum calculations versus group contributions).

Table 4. *Different versions of the PR-EoS used to predict the solvation Gibbs energy of closed-shell molecules.*

Model	**$a_{c,i}$ and b_i**	**α-function**	**c_i**	**σ-profile**
V1 (*tc*-PR EoS + Quantum-based σ-profiles)	experimental values of $T_{c,i}$ and $P_{c,i}$ from DIPPR	Twu 91 α-function using optimized values of L, M, and N from [34]	Optimized values from [34]	Quantum-based σ-profiles extracted from the COSMOTherm database (TZVPD-FINE)
V2 (*tc*-PR EoS + GC-based σ-profiles)	experimental values of $T_{c,i}$ and $P_{c,i}$ from DIPPR	Twu 91 α-function using optimized values of L, M, and N from [34]	Optimized values from [34]	GC-based σ-profiles calculated by equation (26)
V3 (PR EoS + GC methods for pure-component parameters + quantum-based σ-profiles)	Predicted by the GC-methods proposed in equations (17) and (18)	Soave α-function with m_i predicted through the GC- method proposed in equations (22) and (23)	Generalized correlations [34,40] using predicted values of $T_{c,i}$, $P_{c,i}$, and ω_i	Quantum-based σ-profiles extracted from the COSMOTherm database (TZVPD-FINE)
V4 (PR EoS + GC methods for pure-component parameters + GC-based σ-profiles)	Predicted by the GC-methods proposed in equations (17) and (18)	Soave α-function with m_i predicted through the GC- method proposed in equations (22) and (23)	Generalized correlations [34,40] using predicted values of $T_{c,i}$, $P_{c,i}$, and ω_i	GC-based σ-profiles calculated by equation (26)

The results obtained for the solvation Gibbs energy of closed-shell molecules are summarized in Table 5. They were classified according to the association behavior of the binary system, i.e., the ability to form hydrogen bonds. Four types of association can be identified for a binary mixture [22]:

1. Binary mixtures in which association does not take place, like mixtures involving hydrocarbons (non-polar and non-associating molecules).
2. Binary mixtures in which only self-association takes place. For example, binary mixture containing alcohols (polar and self-associating molecules) and hydrocarbons. A self-associating molecule is one that contains both hydrogen donor and acceptor sites.
3. Binary mixtures in which only cross-association takes place. They are mixtures of hydrogen-donor and hydrogen-acceptor molecules, like the one of chloroform and acetone.
4. Binary mixtures in which cross and self-association take place. Any mixtures containing associating molecules, with one of them being a self-associating.

Table 5. *Deviation results for the prediction of solvation Gibbs energy data of infinitely diluted closed-shell molecules considering all versions of the model introduced in* Table 4. *The dataset comprises 65120 experimental points extracted from the CompSol database* [16]. *The results without parentheses correspond to the mean unsigned error (MUE), whereas the results in parentheses correspond to the standard deviation (SD), both in kcal/mol.*

Type of association	**Dataset**	**Model version**			
		V1	**V2**	**V3**	**V4**
1. Non-associating mixtures	32214	0.19 (0.32)	0.24 (0.40)	0.41 (0.56)	0.45 (0.59)
2. Self-association only	15642	0.39 (0.57)	0.65 (0.84)	0.61 (0.77)	0.84 (1.01)
3. Cross-association only	1050	0.39 (0.53)	0.37 (0.55)	0.41 (0.55)	0.39 (0.58)
4. Cross and self-association	16214	0.37 (0.58)	0.43 (0.64)	0.51 (0.74)	0.56 (0.82)
All data	65120	0.28 (0.37)	0.39 (0.45)	0.48 (0.46)	0.57 (0.52)

As expected, the semi-predictive version V1 is the most accurate one since it does not use any input obtained from the group contribution methods. Moreover, the universal constants of COSMO-RS have been fitted based on this version of the model [26], which confirms its better performance. Consequently, the versions based on inputs calculated by group contribution method (especially the version V4, in which the EoS parameters and σ-profiles are obtained by GC), provide predictions with greater deviations with respect to the experimental data considered. Caution should be exercised for mixtures in which only self-association takes place, especially for the version of the model that relies entirely on GC-inputs. In any case, the proposed versions of the EoS are suitable to be extended to free radicals since they are able to give satisfactory predictions for the solvation Gibbs energies of closed-shell solutes.

4.2 GC-methods applied to C/H/O free radicals

4.2.1 Prediction of σ-profiles of free radicals

Before discussing the modelling results of solvation Gibbs energies, it is noteworthy to take a look at the free radical σ-profiles obtained through the GC-based correction method (see Eq. (33)). Some comparisons between σ-profiles calculated from quantum chemistry and from the GC-based correction method are proposed in Figure 8.

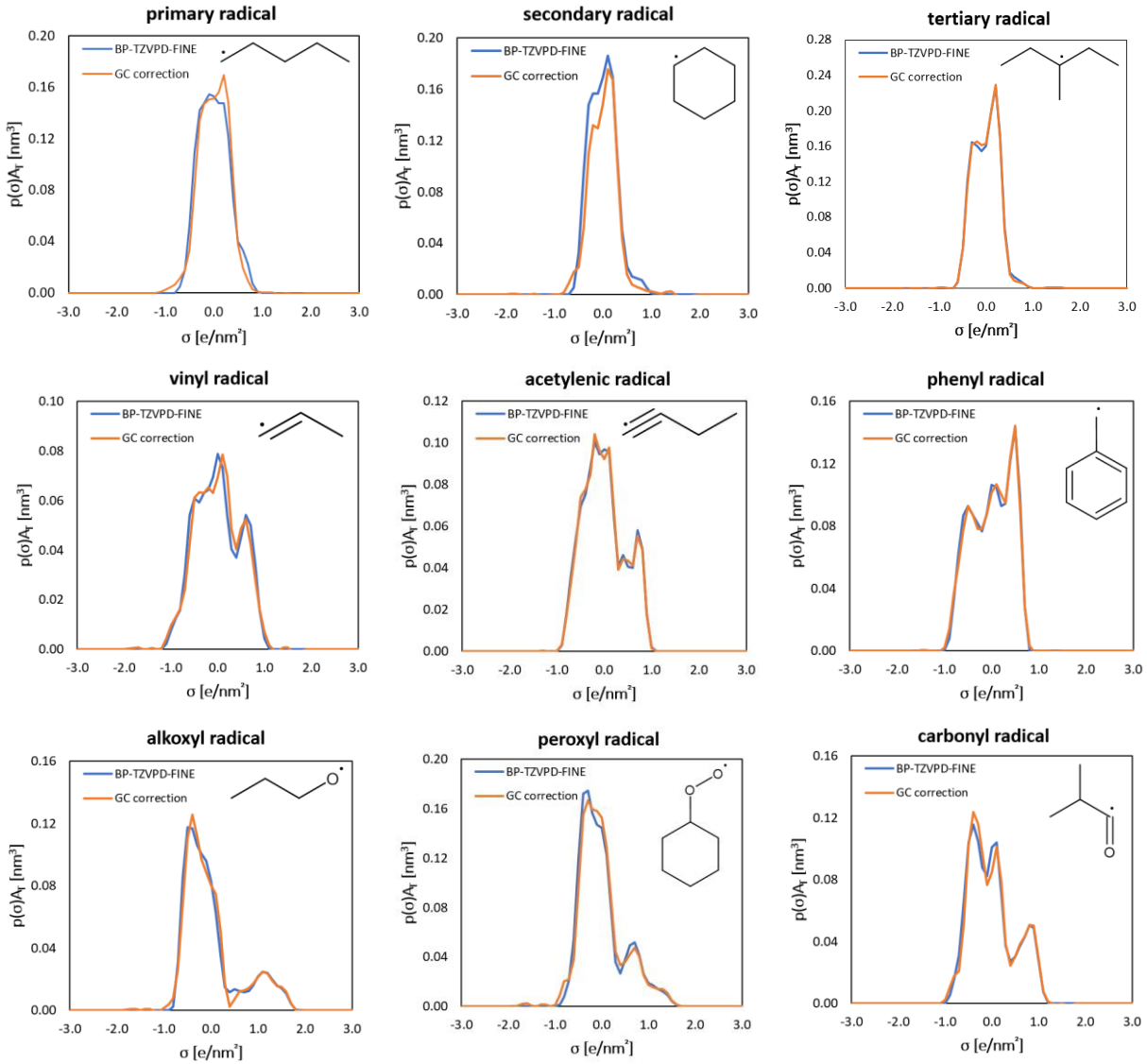


Figure 8. *Examples of free radical σ-profiles predicted by the GC-based correction method (equation (33)). The blue lines are the reference σ-profiles of the radicals obtained through DFT and COSMO calculations at the BP-TZVPD-FINE level of theory. The orange lines are the σ-profiles obtained by applying the GC correction to the quantum chemical σ-profiles of parent molecules.*

As it can be seen in Figure 8, good predictions of σ-profiles can be obtained using the GC-based correction method. The most significant deviations arise in the non-polar region, for charge density values (σ) between -1.0 and $+1.0\ \mathrm{e/nm^2}$, which has a less pronounced impact on the calculation of activity coefficients than in regions of high values of screening charge (for both the positive or negative side of the histogram).

As stated below Eq. (33), it is important to recall that, for the prediction of the free radical σ-profiles in Figure 8, quantum-calculated σ-profiles of the parent molecules were considered. The same could be done with σ-profiles obtained with the group contribution concept, which would lead to similar results, but with larger deviations than in the previous case.

4.2.2 Effect of conformers on the prediction of solvation Gibbs energies

As a first analysis, we will focus on investigating whether relying solely on the most stable conformer compromises the accuracy of *COSMOTherm* in predicting solvation Gibbs energies. In Table 6, we present evaluations for two cases: the first one incorporates all conformers into the calculation, while the second considers only the most stable ones (for both solute and solvent molecules). Following the same classification of binary mixtures used above (see Table 5), the results are categorized based on their association behavior. Again, experimental data extracted from the COMPSOL database were used.

Table 6. *Prediction of experimental solvation Gibbs energy data using COSMOTherm. Effect of using conformers in the accuracy of the model in terms of mean unsigned errors (MUE). The confomers were taken from the cosmo files base provided with the COSMOTherm software.*

Type of association	**Dataset**	**MUE (kcal/mol)**	
		All conformers	**Most stable conformer only**
1. Non-associating mixtures	25979	0.23	0.28
2. Self-association only	17229	0.33	0.37
3. Cross-association only	1206	0.23	0.24
4. Cross and self-association	23384	0.33	0.37
All data	67798*	0.29	0.33

*This dataset is slightly bigger than in the Table 5, because in the previous case only molecules with known decomposition into functional groups were considered.

It was found that for most of the binary systems, employing only the most stable conformer leads to a slight decrease in the model accuracy. However, it is noteworthy that a good level of precision is consistently maintained, with a mean unsigned error of only 0.33 kcal/mol, against 0.29 kcal/mol when all conformers are considered. This analysis is significant because it validates the data generated by *COSMOTherm* on the solvation Gibbs energy of free radicals. Note that this database was constructed based only on the most stable conformation of free radicals, keeping all the solvent conformations, since they were available in the *COSMOTherm* database.

4.2.3 Prediction of solvation Gibbs energies of free radicals

Turning now to the prediction of solvation energies of C/H/O free radicals, Figure 9 details the results obtained for each type of free radical considered in this study. In this base case, the semi-predictive version of the model is considered, relying on equations (30) to (33) for the calculation of the EoS pure parameters and σ-profiles of free radicals. Note that these equations are based on experimental values of $T_{c,i}$, $P_{c,i}$, and ω_i , along with quantum-based σ-profiles of parent molecules (GC-based correction methods are used).

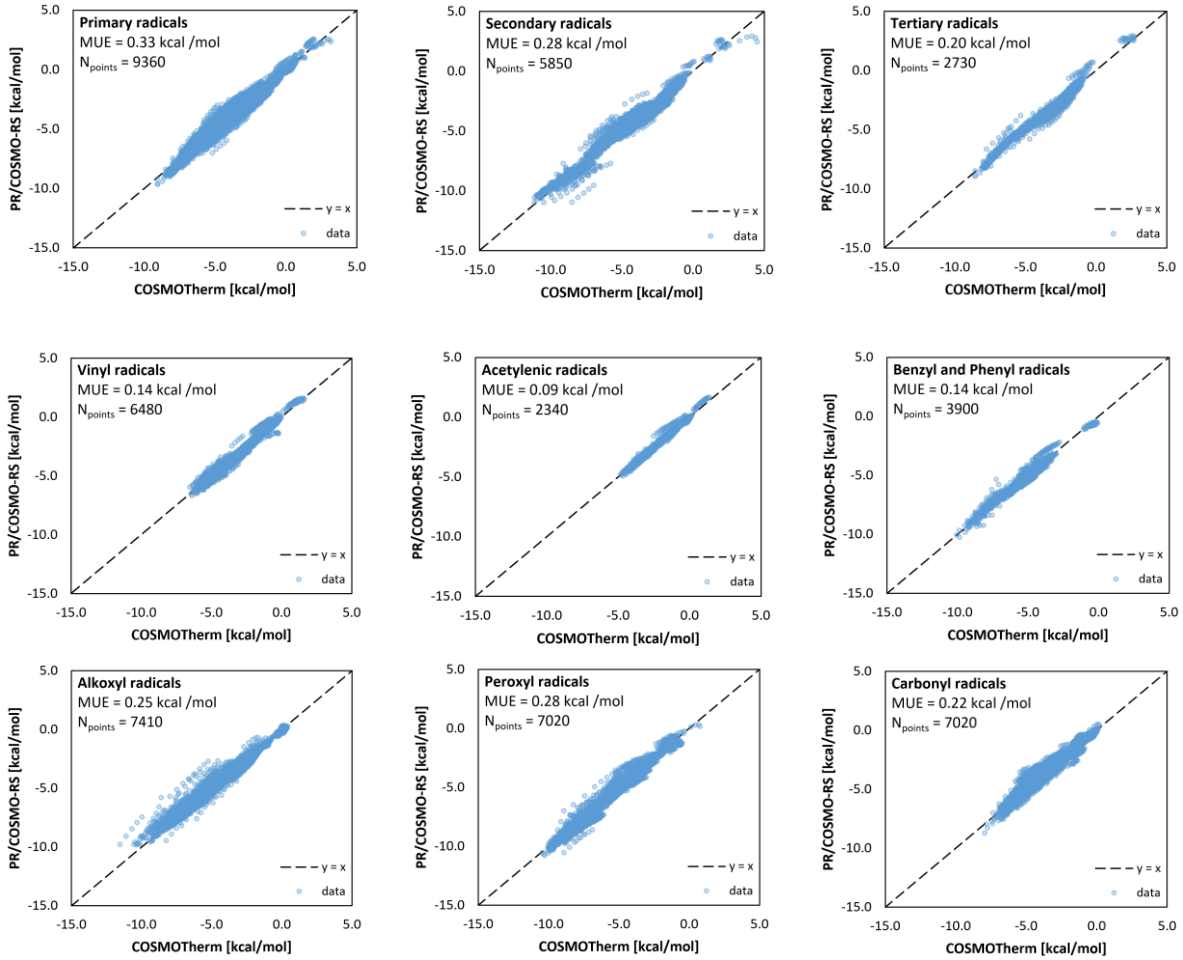


Figure 9. *Parity plots for the solvation Gibbs energy of different types of free radicals in a set of 80 solvents for temperatures between 298 and 398 K. The EoS version V1 was used (see Table 4). The σ-profiles of free radicals were obtained from the GC-correction applied to the quantum-calculated σ-profiles of parent molecules.*

In general, the predictions of the solvation Gibbs energy performed with the proposed model showed a good correlation with the pseudo-experimental database generated for free radicals using the *COSMOTherm* software. The mean unsigned error (MUE) obtained for the entire dataset was about 0.3 kcal/mol. However, when a fully predictive version, entirely based on group contribution methods, is used to predict the EoS parameters and σ-profiles, the deviations increase, as shown in Table 7. This table compares the mean unsigned errors and standard deviations between the two versions of the model.

Table 7. *Deviation results for the prediction of solvation Gibbs energy of infinitely diluted free radicals in a set of 84 solvents. Two versions of the model are used: one based on experimental data for pure component EoS parameters and quantum chemical data for the σ-profile (model version V1, the same used to generate the results presented in Figure 9) and the other based entirely on group contribution methods (model version V4). The results without parentheses correspond to the mean unsigned error (MUE), whereas the results in parentheses correspond to the standard deviation (SD), both in kcal/mol.*

Radical type	**Dataset**	**Inputs for the parent molecule on which the free radical corrections are applied**	
		Exp-input ($T_{c,i}$, $P_{c,i}$, ω_i) QM σ-profile	**GC-input ($a_{c,i}$,b_i, m_i) GC σ-profile**
Primary	9360	0.33 (0.26)	0.56 (0.57)
Secondary	5850	0.28 (0.26)	0.53 (0.43)
Tertiary	2730	0.20 (0.20)	0.42 (0.40)
Vinyl	6480	0.14 (0.15)	0.37 (0.28)
Acetylenic	2440	0.09 (0.11)	0.24 (0.19)
Benzyl and Phenyl	3900	0.14 (0.16)	0.52 (0.34)
Alkoxyl	7410	0.25 (0.25)	0.35 (0.37)
Peroxyl	7020	0.28 (0.25)	0.65 (0.59)
Carbonyl	7020	0.22 (0.25)	0.39 (0.49)
All data	52260	0.23 (0.24)	0.47 (0.47)

Note that the values of the universal constants of COSMO-RS as well as the correction parameters for free radicals have been estimated to get optimal results with the semi-predictive version V1 of the proposed EoS. Naturally, the model reaches its maximum accuracy for this case. The COSMO-RS universal constants as well as the correction parameters for free radicals determined for version V1 are retained in the fully predictive version V4. Using the version V4 entirely based on group contribution for both solvent and solute represents the worst-case scenario in which no experimental and quantum-based input is available. This explains the increase in the mean unsigned errors presented in Table 7 for such a fully predictive version of the model. Nevertheless, this version stands out for its versatility in acquiring all the required inputs through rapid group contribution calculations.

It is important to bear in mind that critical constants and acentric factor ($T_{c,i}$, $P_{c,i}$, and ω_i) are experimentally accessible for the majority of solvents, and their σ-profiles can readily be sourced from databases or generated from DFT calculations. Consequently, the challenge frequently revolves around acquiring the necessary data for the solute molecules. A large range of closed-shell solutes benefits from the availability of $T_{c,i}$, $P_{c,i}$, ω_i, and σ-profiles. Thus, the proposed GC-based methods serve as valuable support for scenarios where such data may not be readily accessible, as is the case with free radical species.

As a result, when the solvent parameters are determined through the semi-predictive model, (i.e., incorporating both experimental and quantum-calculated inputs), and the solute parameters are obtained by the GC methods within the fully predictive version, the resulting deviations generally fall within the range of values shown in Table 7.

4.2.4 *How the solvation Gibbs energy of C/H/O free radicals varies with respect to their parent molecules?*

As mentioned in the introduction, previous kinetic studies dealing with the oxidation of n-alkanes in the liquid phase have assumed that the solvation Gibbs energy of free radicals is exactly the same as that of their parent molecules. It now remains to assess the applicability of this assumption. To this end, Figure 10 evaluates the variation of the solvation Gibbs energy between the different types of free radicals and their parent molecules in two opposite types of solvents: non-associating solvents (non-polar solvents in which there is no hydrogen bonding at all, represented here by n-alkanes) and self-associating solvents (polar solvents in which hydrogen bonding takes place with associating solutes, represented here by n-alcohols).

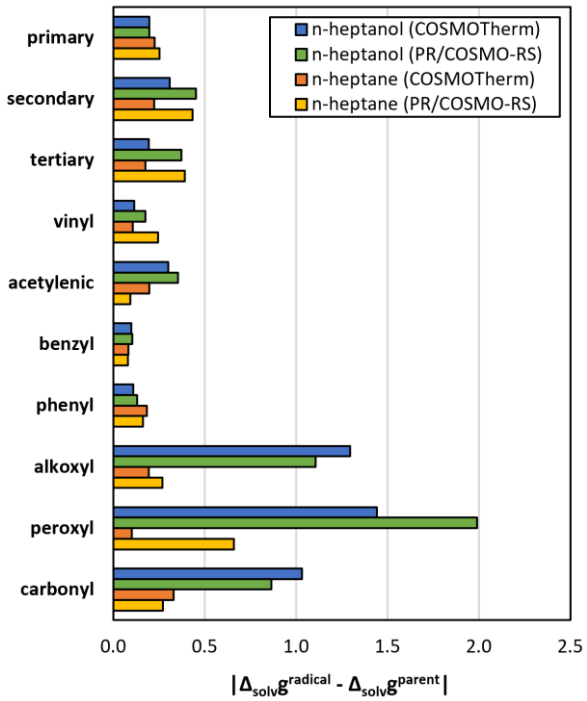


Figure 10. *Average absolute difference between the solvation Gibbs energy of free radicals and their parent molecules in n-alkanes and n-alcohols.*

It is found that in n-alkanes, all radicals have solvation Gibbs energies close to their parent molecules. However, in self-associating solvents such as n-alcohols, the alkoxyl, peroxyl and carbonyl radicals show significant changes in the solvation behavior when compared to their parent molecules. The explanation for this is that these radicals have a reduced hydrogen bonding network due to the loss of the *H* atom from a hydrogen donor site. Since hydrogen bonding stabilizes the solute in the liquid phase, the fact that the free radicals have fewer hydrogen bonding sites than their parent molecule leads to higher values of $\Delta_{\mathrm{solv}}\bar{g}_{liq}$ (in most cases is up to 1.0 kcal/mol higher).

To illustrate this, let us take the molecule 1-propanol, as an example, and then form two different radicals from it: a primary alkyl radical obtained by removing an *H* atom from the CH_3 group, and an alkoxyl radical obtained by removing an *H* atom from the *OH* group. We then solvate these molecules in a self-associating solvent such as ethanol. The results are shown in Figure 11.

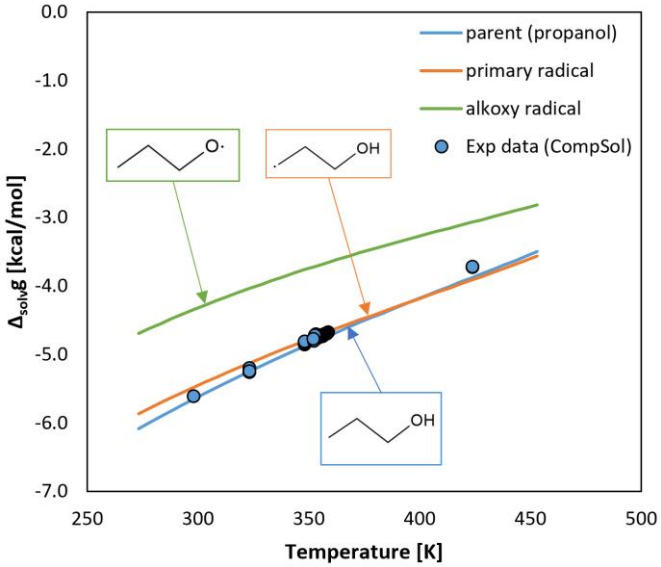


Figure 11. *Solvation Gibbs energy for n-propanol (blue line) and its primary (orange line) and alkoxyl (green line) radicals in ethanol calculated with tc-PR/COSMO-RS model. The blue dots represent real experimental data for the n-propanol/ethanol system extracted from the CompSol database* [16].

It can be seen that the solvation energy of the alkoxy radical is significantly lower (in absolute value) than that of its parent molecule due to the reduction in the number of hydrogen bonds. In such cases, it is not advisable to assume that the solvation energy of the radical is the same as that of the parent molecule.

From the COSMO theory standpoint, the loss of the hydrogen donor site is directly reflected in the σ-profile obtained, as shown in Figure 12. It can be seen that the hydrogen bond donor region of the alkoxy radical completely disappears, and a significant transformation occurs in the hydrogen bond acceptor region. Conversely, the σ-profile of the primary radical undergoes comparatively minor changes in the non-polar region, which, in turn, has a less pronounced influence on the calculation of interaction energies and activity coefficients within the COSMO-RS framework.

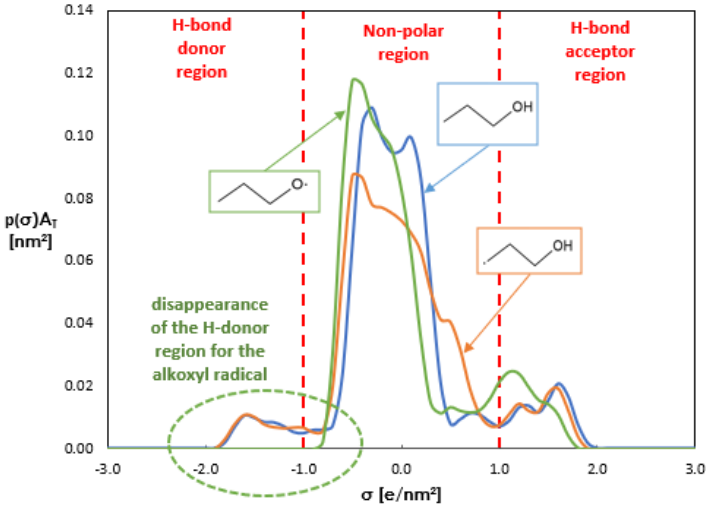


Figure 12. *BP-TZVPD-FINE σ-profiles of 1-propanol (blue line), its primary (orange line) and its alkoxyl radical (green line).*

4.3 Solvation Gibbs energy of C/H/O free radicals in mixtures

4.3.1 Surrogate fuel blends

To discuss the solvation of free radicals in mixtures, blends of fuels and biofuels were chosen as study cases. The first example is a SP95E10 fuel used in spark ignition (Otto cycle) engines. To simplify the problem, the SP95E10 fuel can be simulated as a RON 95 surrogate fuel. Note that RON stands for *Research Octane Number*, which means that a RON 95 fuel contains 95% iso-octane and 5% n-heptane (volume basis). The SP95E10 fuel also contains 10% of oxygenated compounds in its composition which is assumed here to be ethanol only. In this study, we fix the proportions of iso-octane and n-heptane according to RON 95, and vary the amount of ethanol.

As far as diesel engines are concerned, the surrogate fuel considered in this study is a mixture of n-dodecane and m-xylene with a composition in of 77 % and 23 % by volume, respectively. To this mixture of hydrocarbons, biodiesel was added at a molar fraction ranging from 0 to 100%. Biodiesel are esters obtained from the transesterification reaction of vegetable oils with short-chain alcohols. In this study methyl-palmitoleate was considered to represent a biodiesel molecule.

4.3.2 Solvation of peroxyl radicals in the surrogate fuel blends

Considering the pivotal role that peroxyl radicals play in the fuel autoxidation mechanism, we have chosen them, along with alkyl radicals, as illustrative examples. In the context of Otto cycle fuel blends (RON 95 + ethanol), Figure 13 provides insights into the solvation of a peroxyl radical derived from 2-heptyl hydroperoxide and a secondary radical derived from n-heptane. Furthermore, the same figure shows the solvation results for a peroxyl radical formed from 2-decyl hydroperoxide and a benzyl radical generated from m-xylene within the diesel surrogate fuel blend (n-dodecane + m-xylene + methyl palmitoleate). A good level of agreement is observed between the predictions generated by the cubic EoS and COSMOTherm, with diesel fuel showing a better agreement between the two methods.

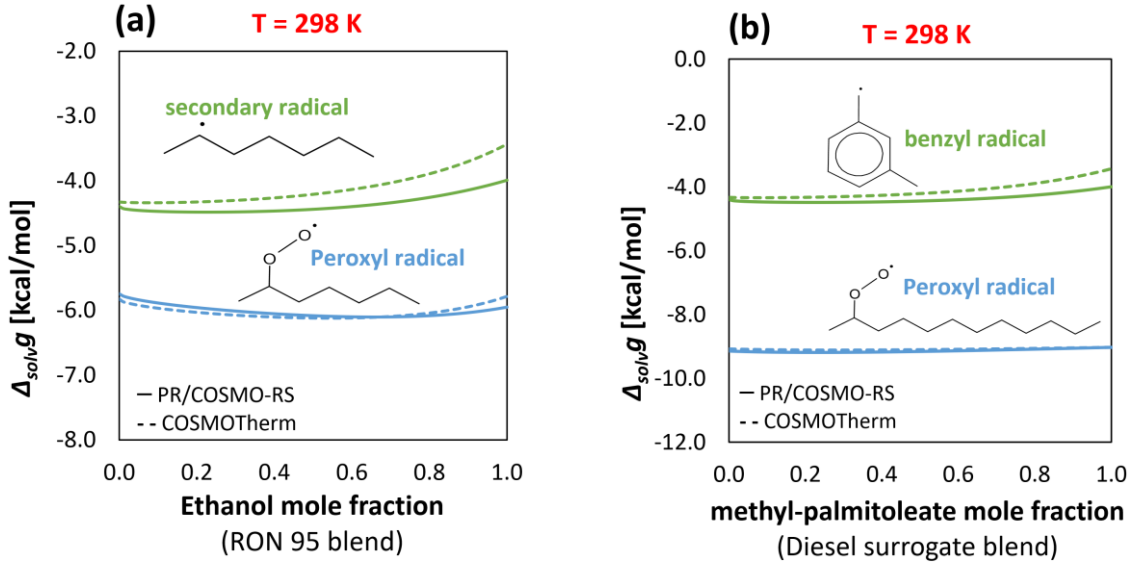


Figure 13. *In (a), the effect the of ethanol concentration in a RON 95 surrogate fuel on the solvation Gibbs energy of a peroxyl and a secondary radical. In (b), the effect of the methyl palmitoleate concentration in a diesel surrogate fuel on the solvation Gibbs of a peroxyl and a benzyl radical. The results were obtained with the proposed EoS (semi-predictive version) and compared with COSMOTherm*

5 Conclusion

The Peng-Robinson (PR) cubic equation of state (EoS) with advanced mixing rules and using COSMO-RS as the activity coefficient (g^E) model was extended to mixtures containing closed-shell molecules and free radicals. Our aim was to propose a model that would allow fast calculation of solvation Gibbs energy of free radicals. This extension was done by means of group contribution-based correction methods developed to predict the PR-EoS parameters ($a_{c,i}$, b_i, and $\alpha_i(T)$) and the COSMO-RS screening charge distribution (σ-profiles) of free radicals from the values of their associated closed-shell molecules.

The accuracy of the proposed model in predicting solvation Gibbs energy data was tested for closed-shell molecules and free radicals. A good correlation between the calculated and reference data was obtained for both cases.

Regarding the change in solvation energy between free radicals and parent molecules, it was observed that hydrogen bonding (or association) plays an important role. Some conclusions can be drawn in this respect:

- In non-associating solvents (such as n-alkanes), the difference in solvation Gibbs energy between free radicals and parent molecules tends to be small, for all types of free radical.
- Alkyl radicals (primary, secondary, tertiary, vinyl, acetylenic, benzyl, and phenyl) have solvation Gibbs energies close to those of their parent molecules, regardless of the nature of the solvent (associating or non-associating).
- The formation of alkoxyl, peroxyl, and carbonyl radicals significantly changes the association behavior. Their solvation Gibbs energy tends to increase significantly in associating solvents compared to their parent molecule.

We believe that this work can be a springboard for the development of flexible kinetic generation frameworks. Our approach, which involves the use of a cubic EoS incorporating COSMO-RS as the g^E model within the mixing rules, has the ability to predict solvation properties for solutes in both liquid and supercritical solvents. This feature distinguishes it from the COSMO-RS model alone, as the latter is applicable to incompressible liquids far from the critical region (see Figure 5 of [26]). This opens the door to a new range of applications, such as supercritical combustion. In addition, the proposed EoS offers the ability to predict solvation properties in mixtures, a task that cannot be achieved by other high-throughput methods such as Abraham's model or machine learning. This valuable feature extends its utility beyond solvent mixtures, allowing compositional changes in any species involved in the mechanism to be accounted for through the calculation of solvation quantities.

However, we would like to draw the reader's attention to the fact that, as it stands, the proposed model does not allow the complete modelling of oxidation mechanisms in the liquid phase, as we have not yet proposed a method for describing the solvation energies of the transition states that determine the value of the activation energy of the process (see equation (3) and Figure 1 in the introduction). Predicting these energies using the methodology presented in this article (i.e. based on a COSMO-type solvation model coupled to a cubic equation of state) is a natural and obvious extension of our work.

Furthermore, we suggest that further research should be undertaken to include radical sites involving other atoms, such as N, S, and P. In a second step, the inclusion of transition state groups can also be performed using the same methodology.

Acknowledgments

The project leading to this application has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 101003318). This project was provided with computer and storage resources by GENCI at IDRIS thanks to the grant 2023-AD010812434R2 on the supercomputer Jean Zay.

Appendix 1: Determination of the exponents β_{ac}, β_b and β_{a0} for the power functions of $a_{c,i}$, b_i, and $a_{0,i}$

It was discussed in section 2.2 that the optimal exponents for the power functions of $a_{c,i}$ and b_i are 0.67 and 0.80, respectively. These values were obtained by minimizing objective functions, as shown in Figure 14. In this figure, the accuracy of the group contribution methods in predicting the constants $a_{c,i}$ and b_i (in terms of the mean percentage error - MPE) is plotted as a function of the value set for the exponents β_{ac} and β_b. It was found that the values 0.67 and 0.80 minimized the MPE obtained.

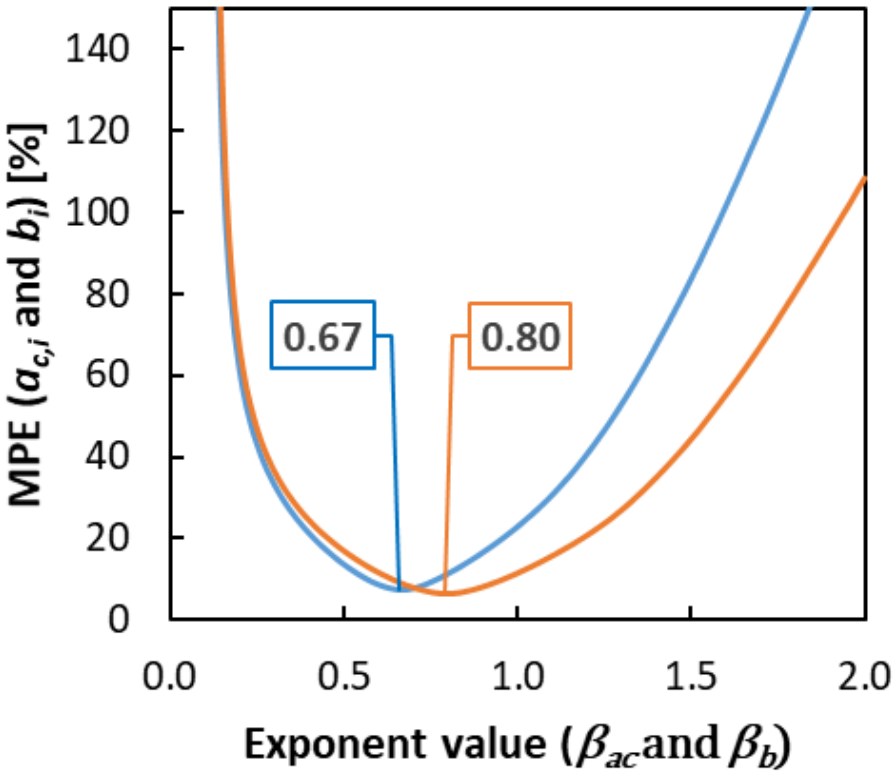


Figure 14. *Plot of the objective functions to be minimized to determine the exponents of the power functions used to calculate $a_{c,i}$ and b_i. MPE is the mean percentage error (MPE) between $a_{c,i}$ and b_i calculated by group contribution methods and using experimental values of critical temperatures and pressures, considering a database of 1613 molecules.*

A further minimization was then performed to determine the value of the β_{a0}, the exponent involved in the GC method for the parameter $a_i(0)$. This latter is used to estimate the Soave α-function shape parameter (m_i). As shown in Figure 15, the optimal value for β_{a0} turns out to be 0.48.

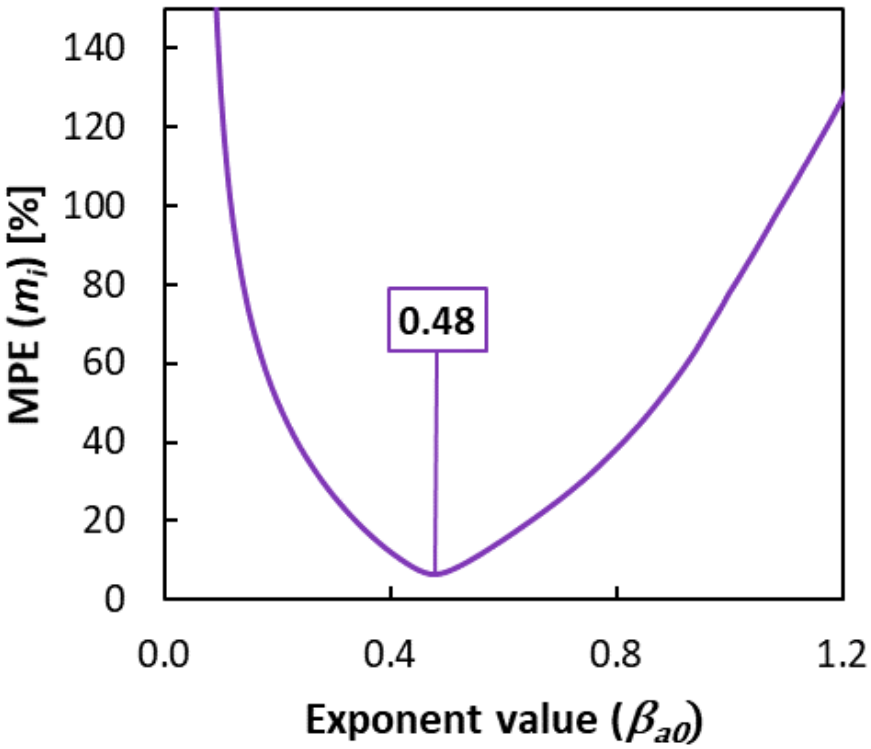


Figure 15. *Plot of the objective functions to be minimized to determine the exponents of the power functions used to calculate m_i. MPE is the mean percentage error (MPE) between m_i calculated by a group contribution method and experimental values of acentric factors, considering a database of 1613 molecules.*

Appendix 2: Functional groups and radical corrections

The UNIFAC groups considered in this study are given in Table 8, while the specifically-defined groups for free radicals are given in Table 9. Note that the parameters associated with each group to be used in the proposed GC methods are given in the supplementary material, together with example decomposition schemes of 1613 molecules and 135 free radicals.

It should be noted that some groups in Table 8 are abbreviations for entire molecules. For example, ACRYL stands for 2-propenenitrile (or acrylonitrile), DMF for dimethylformamide, DOH for glycol, NMP for n-methyl-2-pyrrolidinone, MORPH for morpholine, and so on. Note also that the notation "AC" corresponds to aromatic carbons, and that "#" corresponds to a triple bond.

Table 8. *List of UNIFAC functional groups.*

#	Group	#	Group	#	Group	#	Group	#	Group
1	[CH_3]	31	[$CHNH_2$]	61	[CH_2SH]	91	[AMH_2]	121	[NH_3]
2	[CH_2]	32	[CH_3NH]	62	[FURFURAL]	92	[$AMHCH_3$]	122	[SO_2]
3	[CH]	33	[CH_2NH]	63	[DOH]	93	[$AMHCH_2$]	123	[C(=O)OOH]
4	[C]	34	[CHNH]	64	[I]	94	[$AM(CH3)_2$]	124	[CH_2OOH]
5	[CH2=CH]	35	[CH_3N]	65	[BR]	95	[$AMCH_3CH_2$]	125	[C-O-OH]
6	[CH=CH]	36	[CH_2N]	66	[CH≡C]	96	[$AM(CH_2)_2$]	126	[CHOOH]
7	[CH_2=C]	37	[$ACNH_2$]	67	[C≡C]	97	[$C_2H_5O_2$]	127	[$CHOOCH_2$]
8	[CH=C]	38	[C_5H_5N]	68	[DMSO]	98	[$C_2H_4O_2$]	128	[CH_3OOC]
9	[C=C]	39	[C_5H_4N]	69	[ACRY]	99	[CH_3S]	129	[CH_2ONO_2]
10	[ACH]	40	[C_5H_3N]	70	[CL-(C=C)]	100	[CH_2S]	130	[$CHONO_2$]
11	[AC]	41	[CH_3CN]	71	[ACF]	101	[CHS]	131	[H_2C=O]
12	[$ACCH_3$]	42	[CH_2CN]	72	[DMF]	102	[MORPH]	132	[CH_3OOH]
13	[$ACCH_2$]	43	[COOH]	73	[CF_3]	103	[C_4H_4S]	133	[CH_2CCO]
14	[ACCH]	44	[HCOOH]	74	[CF_2]	104	[C_4H_3S]	134	[H_2O_2]
15	[OH]	45	[CH_2CL]	75	[CF]	105	[C_4H_2S]		
16	[CH3OH]	46	[CHCL]	76	[COO]	106	[H_2]		
17	[H_2O]	47	[CCL]	77	[SIH_3]	107	[O_2]		
18	[ACOH]	48	[CH_2CL_2]	78	[SIH_2]	108	[N_2]		
19	[CH_3CO]	49	[$CHCL_2$]	79	[SIH]	109	[CO]		
20	[CH_2CO]	50	[CCL_2]	80	[SI]	110	[CO_2]		
21	[CH=O]	51	[$CHCL_3$]	81	[SIO]	111	[H_2S]		
22	[CH_3COO]	52	[CCL_3]	82	[NMP]	112	[CH_4]		
23	[CH_2COO]	53	[CCL_4]	83	[CCL_3F]	113	[C_2H_2]		
24	[HCOO]	54	[ACCL]	84	[CCL2F]	114	[C_2H_4]		
25	[CH_3O]	55	[CH_3NO_2]	85	[$HCCL_2F$]	115	[C_2H_6]		
26	[CH_2O]	56	[CH_2NO_2]	86	[HCCLF]	116	[C_3H_6]		
27	[CHO]	57	[$CHNO_2$]	87	[$CCLF_2$]	117	[C_3H_8]		
28	[THF]	58	[$ACNO_2$]	88	[$HCCLF_2$]	118	[C_4H_{10}]		
29	[CH_3NH_2]	59	[CS_2]	89	[$CCLF_3$]	119	[Ar]		
30	[CH_2NH_2]	60	[CH_3SH]	90	[CCL_2F_2]	120	[Cl_2]		

Table 9. *List of specific groups for free radicals. Note that groups that represent entire molecules are indicated by name in the index. The secondary group "33" is different from group "31" because in 33 resonance can occur, changing the position of the radical site, as explained below.*

#	Correction group	#	Correction group
H-atom		*Phenyl*	
1	$[H^{\bullet}(-)H_2]$	23	$[AC^{\bullet}(-)AH]$
Acetylenic		*Primary*	
2	$[C^{\bullet} \equiv C(-)CH \equiv C]$	24	$[C^{\bullet}H_2(-)CH_3]$
3	$[H_2C^{\bullet}(-)CH_2]$	25	$[C_4^{:}H_9(-)C_4H_{10}]_{\text{butane-primary}}$
Alkoxyl		26	$[C_2^{:}H_5(-)C_2H_6]_{\text{ethane-primary}}$
4	$[O^{\bullet}(-)OH]$ water	27	$[C^{\bullet}H_3(-)CH_4]$ methane-primary
5	$[HO^{\bullet}(-)H_2O]$	28	$[C^{\bullet}H_3OH(-)CH_3OH]$ methanol-primary
6	$[CH_3O^{\bullet}(-)CH_3OH]$ methanol	29	$[C_3^{:}H_7(-)C_3H_8]_{\text{propane-primary}}$
7	$[ACO^{\bullet}(-)ACOH]$	30	$[C_3^{:}H_5(-)C_3H_6]_{\text{propene-primary}}$
8	$[ACO^{\bullet}(-)ACOH]_{\text{phenol}}$	*Secondary*	
Benzyl		31	$[C^{\bullet}H(-)CH_2]$
9	$[ACC^{\bullet}H_2(-)ACCH_3]$	32	$[C_4^{:}H_9(-)C_4H_{10}]_{\text{butane-secondary}}$
10	$[ACC^{\bullet}H(-)ACCH_2]$	33	$[C^{\bullet}H(-)CH_2]_{\text{resonance}}$
Carbonyl and carboxyl		34	$[C_3^{:}H_7(-)C_3H_8]_{\text{propane-secondary}}$
11	$[C^{\bullet} = O(-)CH = O]$	*Tertiary*	
12	$[COO^{\bullet}(-)COOH]$	35	$[C^{\bullet}(-)CH]$
13	$[C^{\bullet}OO(-)HCOO]$	*Vinyl*	
14	$[HC^{\bullet} = O(-)HCH = O]$	36	$[C^{\bullet} = C(-)CH = C]$
15	$[HCOO^{\bullet}(-)HCOOH]$ formicacid	37	$[CH^{\bullet} = C(-)CH_2 = C]$
16	$[C^{\bullet}OOH(-)HCOOH]$ formicacid	38	$[CH^{\bullet} = CH_2(-)CH_2 = CH_2]$ ethene-vinyl
Peroxyl		39	$[CH^{\bullet}CC = O(-)CH_2CC = O]$ ethenone-vinyl
17	$[HO_2^{:}(-)H_2O_2]$ hydrogen-peroxide	40	$[C^{\bullet} = CH(-)CH = CH]$
18	$[C(= O)OO^{\bullet}(-)C(= O)OOH]$	41	$[CH_2 = C^{\bullet}(-)CH_2 = CH]$
19	$[CHOO^{\bullet}(-)CHOOH]$	42	$[C_3^{:}H_5(-)C_3H_6]_{\text{propene-vinyl}}$
20	$[CH_3OO^{\bullet}(-)CH_3OOH]$ methylhydroperoxide		
21	$[CH_2OO^{\bullet}(-)CH_2OOH]$		
22	$[C - O - O^{\bullet}(-)C - O - OH]$		

In Table 9, some of groups are also whole small molecules. This was done because the first molecules in a homologous series with a small molecular weight tend to deviate from the trend of the group contribution method. In this sense, the inclusion of these groups improves the accuracy of the method. Note also that groups in which the unpaired electron is delocalized due to resonance effects have also been treated separately. For example, an alkyl radical formed near a carbonyl is actually either alkyl (secondary) or alkoxyl radical, as indicated in the figure below.

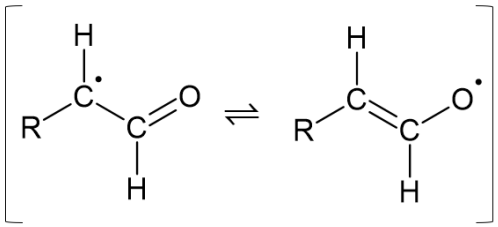


Figure 16. *Resonance hybrid of a secondary radical formed in the α position of a carbonyl (C=O). Note that this radical is either alkyl (secondary) or alkoxyl.*

References

[1] L.P. de Oliveira, D. Hudebine, D. Guillaume, J.J. Verstraete, A Review of Kinetic Modeling Methodologies for Complex Processes, Oil Gas Sci. Technol. – Rev. D'IFP Energ. Nouv. 71 (2016) 45. https://doi.org/10.2516/ogst/2016011.

[2] S.W. Benson, F.R. Cruickshank, D.M. Golden, G.R. Haugen, H.E. O'Neal, A.S. Rodgers, R. Shaw, R. Walsh, Additivity rules for the estimation of thermochemical properties, Chem. Rev. 69 (1969) 279–324. https://doi.org/10.1021/cr60259a002.

[3] R. Van de Vijver, N.M. Vandewiele, P.L. Bhoorasingh, B.L. Slakman, F. Seyedzadeh Khanshan, H.-H. Carstensen, M.-F. Reyniers, G.B. Marin, R.H. West, K.M. Van Geem, Automatic Mechanism and Kinetic Model Generation for Gas- and Solution-Phase Processes: A Perspective on Best Practices, Recent Advances, and Future Challenges: AUTOMATIC MECHANISM GENERATION FOR GAS- AND SOLUTION-PHASE PROCESSES, Int. J. Chem. Kinet. 47 (2015) 199–231. https://doi.org/10.1002/kin.20902.

[4] F. Battin-Leclerc, E. Blurock, R. Bounaceur, R. Fournet, P.-A. Glaude, O. Herbinet, B. Sirjean, V. Warth, Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models, Chem. Soc. Rev. 40 (2011) 4762. https://doi.org/10.1039/c0cs00207k.

[5] Kh.E. Kharlampidi, T.Sh. Nurmurodov, N.V. Ulitin, K.A. Tereshchenko, N.P. Miroshkin, D.A. Shiyan, N.A. Novikov, O.V. Stoyanov, N.N. Ziyatdinov, T.V. Lapteva, S.L. Khursan, Design of cumene oxidation process, Chem. Eng. Process. - Process Intensif. 161 (2021) 108314. https://doi.org/10.1016/j.cep.2021.108314.

[6] M. Weber, J.G. Daldrup, M. Weber, Noncatalyzed Radical Chain Oxidation: Cumene Hydroperoxide, in: S.S. Stahl, P.L. Alsters (Eds.), Liq. Phase Aerob. Oxid. Catal. Ind. Appl. Acad. Perspect., 1st ed., Wiley, 2016: pp. 15–31. https://doi.org/10.1002/9783527690121.ch2.

[7] A. Vomeri, M. Stucchi, A. Villa, C. Evangelisti, A. Beck, L. Prati, New insights for the catalytic oxidation of cyclohexane to K-A oil, J. Energy Chem. 70 (2022) 45–51. https://doi.org/10.1016/j.jechem.2022.02.008.

[8] A. Abutaleb, M.A. Ali, A comprehensive and updated review of studies on the oxidation of cyclohexane to produce ketone-alcohol (KA) oil, Rev. Chem. Eng. 38 (2022) 769–797. https://doi.org/10.1515/revce-2020-0059.

[9] S.H. Lin, C.C. Lo, Fenton process for treatment of desizing wastewater, Water Res. 31 (1997) 2050–2056. https://doi.org/10.1016/S0043-1354(97)00024-9.

[10] E.G. Jones, L.M. Balster, W.J. Balster, Autoxidation of Aviation Fuels in Heated Tubes: Surface Effects, Energy Fuels 10 (1996) 831–836. https://doi.org/10.1021/ef9502012.

[11] M.D. Le, V. Warth, L. Giarracca, E. Moine, R. Bounaceur, R. Privat, J.-N. Jaubert, R. Fournet, P.-A. Glaude, B. Sirjean, Development of a detailed kinetic model for the oxidation of *n* -butane in the liquid phase, J. Phys. Chem. B 125 (2021) 6955–6967. https://doi.org/10.1021/acs.jpcb.1c02988.

[12] M. Alves-Fortunato, E. Ayoub, K. Bacha, A. Mouret, C. Dalmazzone, Fatty Acids Methyl Esters (FAME) autoxidation: New insights on insoluble deposit formation process in biofuels, Fuel 268 (2020) 117074. https://doi.org/10.1016/j.fuel.2020.117074.

[13] A. Jalan, R.H. West, W.H. Green, An Extensible Framework for Capturing Solvent Effects in Computer Generated Kinetic Models, J. Phys. Chem. B 117 (2013) 2955–2970. https://doi.org/10.1021/jp310824h.

[14] J.T. Edward, Molecular volumes and the Stokes-Einstein equation, J. Chem. Educ. 47 (1970) 261. https://doi.org/10.1021/ed047p261.

[15] A. Ben-Naim, Solvation Thermodynamics,.

[16] E. Moine, R. Privat, B. Sirjean, J.-N. Jaubert, Estimation of solvation quantities from experimental thermodynamic data: development of the comprehensive CompSol databank for pure and mixed solutes, J. Phys. Chem. Ref. Data 46 (2017) 033102. https://doi.org/10.1063/1.5000910.

[17] Y. Chung, F.H. Vermeire, H. Wu, P.J. Walker, M.H. Abraham, W.H. Green, Group Contribution and Machine Learning Approaches to Predict Abraham Solute Parameters, Solvation Free Energy, and Solvation Enthalpy, J. Chem. Inf. Model. 62 (2022) 433–446. https://doi.org/10.1021/acs.jcim.1c01103.

[18] H. Lim, Y. Jung, MLSolvA: solvation free energy prediction from pairwise atomistic interactions by machine learning, J. Cheminformatics 13 (2021) 56. https://doi.org/10.1186/s13321-021-00533-z.

[19] F.H. Vermeire, W.H. Green, Transfer learning for solvation free energies: From quantum chemistry to experiments, Chem. Eng. J. 418 (2021) 129307. https://doi.org/10.1016/j.cej.2021.129307.

[20] A. Jalan, R.W. Ashcraft, R.H. West, W.H. Green, Predicting solvation energies for kinetic modeling, Annu. Rep. Sect. C Phys. Chem. 106 (2010) 211. https://doi.org/10.1039/b811056p.

[21] A. Klamt, M. Diedenhofen, Calculation of Solvation Free Energies with DCOSMO-RS, J. Phys. Chem. A 119 (2015) 5439–5445. https://doi.org/10.1021/jp511158y.

[22] F.C. Paes, R. Privat, J.-N. Jaubert, B. Sirjean, A comparative study of COSMO-based and equation-of-state approaches for the prediction of solvation energies based on the compsol databank, Fluid Phase Equilibria 561 (2022) 113540. https://doi.org/10.1016/j.fluid.2022.113540.

[23] E. Moine, R. Privat, J.-N. Jaubert, B. Sirjean, N. Novak, E. Voutsas, C. Boukouvalas, Can we safely predict solvation Gibbs energies of pure and mixed solutes with a cubic equation of state?, Pure Appl. Chem. 91 (2019) 1295–1307. https://doi.org/10.1515/pac-2018-1112.

[24] E. Voutsas, K. Magoulas, D. Tassios, Universal Mixing Rule for Cubic Equations of State Applicable to Symmetric and Asymmetric Systems: Results with the Peng−Robinson Equation of State, Ind. Eng. Chem. Res. 43 (2004) 6238–6246. https://doi.org/10.1021/ie049580p.

[25] M.D. Le, Z. El Sayah, R. Benrabah, V. Warth, P.-A. Glaude, R. Privat, R. Fournet, B. Sirjean, An experimental and detailed kinetic modeling of the thermal oxidation stability of *n*-decane as a jet fuel surrogate component, Fuel 342 (2023) 127754. https://doi.org/10.1016/j.fuel.2023.127754.

[26] F.C. Paes, R. Privat, J.-N. Jaubert, B. Sirjean, Prediction of solvation energies at infinite dilution by the *tc*-PR cubic equation of state with advanced mixing rule based on COSMO-RS as g^E model, J. Mol. Liq. 386 (2023) 122480. https://doi.org/10.1016/j.molliq.2023.122480.

[27] D. Constantinescu, A. Klamt, D. Geană, Vapor–liquid equilibrium prediction at high pressures using activity coefficients at infinite dilution from COSMO-type methods, Fluid Phase Equilibria 231 (2005) 231–238. https://doi.org/10.1016/j.fluid.2005.01.014.

[28] Y. Shimoyama, Y. Iwai, S. Takada, Y. Arai, T. Tsuji, T. Hiaki, Prediction of phase equilibria for mixtures containing water, hydrocarbons and alcohols at high temperatures and pressures by cubic equation of state with GE type mixing rule based on COSMO-

RS, Fluid Phase Equilibria 243 (2006) 183–192. https://doi.org/10.1016/j.fluid.2006.03.007.

[29] K. Leonhard, J. Veverka, K. Lucas, A comparison of mixing rules for the combination of COSMO-RS and the Peng–Robinson equation of state, Fluid Phase Equilibria 275 (2009) 105–115. https://doi.org/10.1016/j.fluid.2008.09.016.

[30] L.F.K. Possani, P.B. Staudt, R. de P. Soares, Prediction of water solubilities in hydrocarbons and oils using F-SAC coupled with SRK–EoS, Fluid Phase Equilibria 427 (2016) 394–405. https://doi.org/10.1016/j.fluid.2016.08.001.

[31] D.-Y. Peng, D.B. Robinson, A new two-constant equation of state, Ind. Eng. Chem. Fundam. 15 (1976) 59–64. https://doi.org/10.1021/i160057a011.

[32] A. Pina-Martinez, R. Privat, S. Lasala, G. Soave, J.-N. Jaubert, Search for the optimal expression of the volumetric dependence of the attractive contribution in cubic equations of state, Fluid Phase Equilibria 522 (2020) 112750. https://doi.org/10.1016/j.fluid.2020.112750.

[33] Y. Le Guennec, R. Privat, J.-N. Jaubert, Development of the translated-consistent *tc*-PR and *tc*-RK cubic equations of state for a safe and accurate prediction of volumetric, energetic and saturation properties of pure compounds in the sub- and super-critical domains, Fluid Phase Equilibria 429 (2016) 301–312. https://doi.org/10.1016/j.fluid.2016.09.003.

[34] A. Pina-Martinez, R. Privat, J. Jaubert, Use of 300,000 pseudo- experimental data over 1800 pure fluids to assess the performance of four cubic equations of state: SRK, PR, *tc*- RK, and *tc*- PR, AIChE J. 68 (2022) e17518. https://doi.org/10.1002/aic.17518.

[35] G. Soave, Equilibrium constants from a modified Redlich-Kwong equation of state, Chem. Eng. Sci. 27 (1972) 1197–1203. https://doi.org/10.1016/0009-2509(72)80096-4.

[36] P.M. Mathias, T.W. Copeman, Extension of the Peng-Robinson equation of state to complex mixtures: Evaluation of the various forms of the local composition concept, Fluid Phase Equilibria 13 (1983) 91–108. https://doi.org/10.1016/0378-3812(83)80084-3.

[37] C.H. Twu, D. Bluck, J.R. Cunningham, J.E. Coon, A cubic equation of state with a new alpha function and a new mixing rule, Fluid Phase Equilibria 69 (1991) 33–50. https://doi.org/10.1016/0378-3812(91)90024-2.

[38] Y. Le Guennec, R. Privat, S. Lasala, J.-N. Jaubert, On the imperative need to use a consistent α-function for the prediction of pure-compound supercritical properties with a cubic equation of state, Fluid Phase Equilibria 445 (2017) 45–53. https://doi.org/10.1016/j.fluid.2017.04.015.

[39] Y. Le Guennec, S. Lasala, R. Privat, J.-N. Jaubert, A consistency test for α-functions of cubic equations of state, Fluid Phase Equilibria 427 (2016) 513–538. https://doi.org/10.1016/j.fluid.2016.07.026.

[40] K. Magoulas, D. Tassios, Thermophysical properties of n-Alkanes from C1 to C20 and their prediction for higher ones, Fluid Phase Equilibria 56 (1990) 119–140. https://doi.org/10.1016/0378-3812(90)85098-U.

[41] M.-J. Huron, J. Vidal, New mixing rules in simple equations of state for representing vapour-liquid equilibria of strongly non-ideal mixtures, Fluid Phase Equilibria 3 (1979) 255–271. https://doi.org/10.1016/0378-3812(79)80001-1.

[42] M.L. Michelsen, A modified Huron-Vidal mixing rule for cubic equations of state, Fluid Phase Equilibria 60 (1990) 213–219. https://doi.org/10.1016/0378-3812(90)85053-D.

[43] A. Pina-Martinez, R. Privat, I.K. Nikolaidis, I.G. Economou, J.-N. Jaubert, What is the optimal activity coefficient model to be combined with the *translated–consistent* Peng–Robinson equation of state through advanced mixing rules? Cross-comparison and grading of the Wilson, UNIQUAC, and NRTL a^E models against a benchmark database involving 200 binary systems, Ind. Eng. Chem. Res. 60 (2021) 17228–17247. https://doi.org/10.1021/acs.iecr.1c03003.

[44] R. Privat, J.-N. Jaubert, Y. Le Guennec, Incorporation of a volume translation in an equation of state for fluid mixtures: which combining rule? Which effect on properties of mixing?, Fluid Phase Equilibria 427 (2016) 414–420. https://doi.org/10.1016/j.fluid.2016.07.035.

[45] E. Mullins, R. Oldland, Y.A. Liu, S. Wang, S.I. Sandler, C.-C. Chen, M. Zwolak, K.C. Seavey, Sigma-Profile Database for Using COSMO-Based Thermodynamic Methods, Ind. Eng. Chem. Res. 45 (2006) 4389–4415. https://doi.org/10.1021/ie060370h.

[46] J.-W. Qian, R. Privat, J.-N. Jaubert, P. Duchet-Suchaux, Enthalpy and heat capacity changes on mixing: fundamental aspects and prediction by means of the PPR78 cubic equation of state, Energy Fuels 27 (2013) 7150–7178. https://doi.org/10.1021/ef401605c.

[47] A. Pina-Martinez, R. Privat, J.-N. Jaubert, D.-Y. Peng, Updated versions of the generalized Soave α-function suitable for the Redlich-Kwong and Peng-Robinson equations of state, Fluid Phase Equilibria 485 (2019) 264–269. https://doi.org/10.1016/j.fluid.2018.12.007.

[48] T. Mu, J. Rarey, J. Gmehling, Group contribution prediction of surface charge density profiles for COSMO-RS(Ol), AIChE J. 53 (2007) 3231–3240. https://doi.org/10.1002/aic.11338.

[49] T. Mu, J. Rarey, J. Gmehling, Group contribution prediction of surface charge density distribution of molecules for COSMO-SAC, AIChE J. 55 (2009) 3298–3300. https://doi.org/10.1002/aic.11933.

[50] C. Hsieh, S. Lin, Determination of cubic equation of state parameters for pure fluids from first principle solvation calculations, AIChE J. 54 (2008) 2174–2181. https://doi.org/10.1002/aic.11552.

[51] A. Klamt, G. Schüürmann, COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, J Chem Soc Perkin Trans 2 (1993) 799–805. https://doi.org/10.1039/P29930000799.

[52] L. Constantinou, R. Gani, New group contribution method for estimating properties of pure compounds, AIChE J. 40 (1994) 1697–1710. https://doi.org/10.1002/aic.690401011.

[53] J. Marrero, R. Gani, Group-contribution based estimation of pure component properties, Fluid Phase Equilibria 183–184 (2001) 183–208. https://doi.org/10.1016/S0378-3812(01)00431-9.

[54] W.H. Carande, A. Kazakov, C. Muzny, M. Frenkel, Quantitative Structure–Property Relationship Predictions of Critical Properties and Acentric Factors for Pure Compounds, J. Chem. Eng. Data 60 (2015) 1377–1387. https://doi.org/10.1021/je501093v.

[55] S. Biswas, Y. Chung, J. Ramirez, H. Wu, W.H. Green, Predicting Critical Properties and Acentric Factors of Fluids Using Multitask Machine Learning, J. Chem. Inf. Model. 63 (2023) 4574–4588. https://doi.org/10.1021/acs.jcim.3c00546.

[56] L. Constantinou, R. Gani, J.P. O'Connell, Estimation of the acentric factor and the liquid molar volume at 298 K using a new group contribution method, Fluid Phase Equilibria 103 (1995) 11–22. https://doi.org/10.1016/0378-3812(94)02593-P.