A New Semi-Empirical Model for Saturated Vapor Density of Pure Compounds

Alexander Kud, Christophe Coquelet, Stefan Maixner

HAL Id: hal-02493875
https://hal.science/hal-02493875

Submitted on 28 Feb 2020

A New Semi-Empirical Model for Saturated Vapor Density of Pure Compounds

Alexander Kud[*], Department of Research, BASF SE, Carl-Bosch-Str. 38, Ludwigshafen 67056, Germany

Christophe Coquelet, Mines ParisTech, PSL University, CTP – Centre Thermodynamics of Processes, 35 rue Saint Honoré, Fontainebleau Cedex 77305, France

Stefan Maixner[**], Department of Research, BASF SE, Carl-Bosch-Str. 38, Ludwigshafen 67056, Germany

ABSTRACT

Many empirical models exist to describe the saturated vapor density. The majority of these models are detached from a physics basis, which often makes it difficult to select the most appropriate model for further thermodynamic modeling. Such empirical models are still used today because many equations of state that do not describe vapor density well enough.

The model presented here is based on a physical and a heuristic equation. It offers many advantages because it includes both the vapor pressure and the compressibility factor information for the saturated vapor line. In this manner, it is possible to describe the vapor density with one equation and only three parameters. The new model is independent of the compounds structure because it is based on vapor density measurements. It was tested in comparison to six frequently used models in terms of accuracy and thermodynamic consistency. The new model received the best rating and provides useful tool for thermodynamic modeling and process design. The assessment of the models was based on Monte Carlo simulations and strict statistical methods for validating models. The new model can also be used in a cubic equation of state for the vapor pressure and vapor density correction.

---

[*] retired 2018. [**] retired 2019.

1 INTRODUCTION

In the past, many working groups have developed correlations to describe the vapor density of pure substances at saturation. These models are not Equations of State (EoS) and are known as heuristic and empirical correlations; they only describe the relationship between the vapor density, the measured variable and the temperature. Similar equations have also been developed for the vapor pressures of pure substances. These equations are based on the principle that $ln(p) = f\left(\frac{1}{T}\right)$. These correlations have the great advantage that they do not require any equations of state to be solved, nor do they require a specific algorithm or a phase stability analysis. In many cases, they are easily parameterized. For some equations, parameter values can be found in Table collections, software database or publications. The new model proposed for vapor density calculation also belongs to this group of heuristic models, such as the vapor pressure equations. Our model is a semi-empirical approach and is compared to the state of the art.

In order to assess the models or models fairly, statistical and physical evaluation quantities were selected with which a so-called H_0 Hypothesis for the model acceptance can be formulated. If this H_0 Hypothesis is not fulfilled, the model is rejected. With these evaluation quantities the models were tested for

- Goodness of Fit and the predictive capability,

- Goodness of estimated model parameter,

- Physical consistency and their predictive capability

  based on a matching H_0 hypothesis.

Since the assessment quantities fluctuate in a statistical sense, Monte Carlo simulations are applied to strengthen the significance of the assessment results obtained. Since the test methods are very extensive and many results are obtained from the investigated models for the individual compounds, the reader should be given a brief overview of the topics reported on:

1. Introduction: Presentation of the state of the art in vapor density description

2. Data base: Measurement data of the compounds

3. Computational Methods

   - Development of the new model

   - Overview of the six test methods

   - Definition of the parameter estimation problem

4. Results and discussion

   - Model assessment based on every test method

   - Summary of the rating

5. Conclusion

Table 1 gives an overview of the models or models for vapor density description discussed in this paper. Models 1 to 7, 11 and 12 represent the state of the art. The new model is represented by model numbers 8 to 10. Table 1 does not claim to be complete because its only purpose is to demonstrate the equation structures.

The state of the art will be presented. To describe the vapor density ρ_v at saturation, Model 1 requires the corresponding liquid density ρ_l of the liquid at saturation. It is only applicable for conditions close to the critical point, and therefore, its applicability is extremely limited. This model is excluded from the evaluation because it does not exhibit any explicit dependence on the absolute temperature T either which is a prerequisite in this work. All other models show a

temperature dependency and are presented as a function of the reduced temperature $\tau = T/T_c$, where the subscript c denotes critical state. Models 2 to 5 in Table 1 are highly similar to each other and applicable across a wide temperature range mostly from the triple point temperature T_{Tp} to the critical temperature. Additional models exist; however, they are similar to those listed in Table 1. Models 2 to 5 and similar models have the following aspects in common:

- They are polynomial models with at least 6 parameters $\Theta_1 \dots \Theta_6$ to be estimated.

- For each compound, the appropriate model for estimating the parameters must be determined. This search can be time consuming.

- They calculate the vapor density extremely accurately within the experimental uncertainty.

Model 6 is based on the description of the liquid and vapor phases. An asymptotic behavior law[1] for the density expansion in the vicinity of the critical point is used and coupled with a rectilinear equation for the coexistence curve[1]. Promising results have been achieved with this model[2]. Only two parameters have to be adjusted. The parametrization of this correlation requires both experimental vapor and liquid densities at saturation. In this study, we have tested the performance of Model 6 in a range of temperatures from the triple point to the critical temperature and we used the eq 5 for vapor density in reference 2 only. If eq 5 in reference 2 is used for the vapor density description alone, the range of values of the estimated parameters and the working range of temperature are increased. To further improve the calculation, we suggest in this work to add one additional parameter, as defined in Model 7. This results in a more flexible correlation than the model considered in Model 6.

Table 1. Overview of the different vapor density models used in this work.

Model n°	Model[a]	N° of parameters	Authors	Comments
1	$\frac{1}{2}(\rho_l + \rho_v) = \rho_c + \Theta_1(\rho_l - \rho_v)^{\frac{10}{3}}$	1	Hakala[9]	only close to the critical point
2	$\rho_v = \Theta_1 + \Theta_2(1-\tau)^{\frac{1}{3}} + \Theta_3(1-\tau)^{\frac{2}{3}} + \Theta_4(1-\tau)$ $+ \Theta_5(1-\tau)^{\frac{4}{3}} + \Theta_6(1-\tau)^{\frac{5}{3}}$	6	Hales[6]	triple point to critical point
3	$\ln\frac{\rho_v}{\rho_c} = \frac{1}{\tau}\left[\Theta_1(1-\tau)^{0.348} + \Theta_2(1-\tau)^{\frac{1}{6}} + \Theta_3(1-\tau)^{\frac{1}{3}} + \Theta_4(1-\tau)^{\frac{2}{3}} + \Theta_5(1-\tau)^{\frac{16}{6}} + \Theta_6(1-\tau)^{\frac{34}{6}}\right]$	6	Guder[3]	triple point to critical point
4	$\ln\frac{\rho_v}{\rho_c} = \frac{1}{\tau}\left[\Theta_1(1-\tau)^{0.346} + \Theta_2(1-\tau)^{\frac{5}{6}} + \Theta_3(1-\tau) + \Theta_4(1-\tau)^2 + \Theta_5(1-\tau)^3 + \Theta_6(1-\tau)^5\right]$	6	Funke[4]	triple point to critical point
5	$\ln\frac{\rho_v}{\rho_c} = \Theta_1(1-\tau)^{\frac{2}{6}} + \Theta_2(1-\tau)^{\frac{4}{6}} + \Theta_3(1-\tau)^{\frac{8}{6}}$ $+ \Theta_4(1-\tau)^{\frac{18}{6}} + \Theta_5(1-\tau)^{\frac{37}{6}} + \Theta_6(1-\tau)^{\frac{71}{6}}$	6	Wagner[5]	triple point to critical point
6	$\rho_v = \rho_c + \Theta_1 T_c(1-\tau) + \frac{\Theta_2 T_c^{\beta}}{2}(1-\tau)^{\beta}$ with $\beta = 0.325$	2	Coquelet et al.[2]	combination of the rectilinear diameter equation of the coexistence curve and the scaling equation in the vicinity of the critical point (expansion for the density) from $\tau = 0.6$ to $\tau = 1$
7	$\rho_v = \rho_c + \Theta_1 T_c(1-\tau) + \frac{\Theta_2 T_c^{\Theta_2}}{2}(1-\tau)^{\Theta_3}$	3	this work	one more parameter than model 6 to obtain a more flexible model from $\tau = 0.6$ to $\tau = 1$
8 to 10	"New model" in this work eqs 2 and 4. In model 8: $Z_c = p_c v_c / \mathcal{R} T_c$	3	this work	triple point to critical point

In Model 9: $Z_c = 1/3$ for SRK[b] EoS[36]

In Model 10: $Z_c = 0.307$ for PR[c] EoS[37]

| 11 | $$Z_r = A(Z_c,\Theta) + B(Z_c,\Theta)\cdot\pi + \frac{C(Z_c,\Theta)\cdot D(Z_c,\Theta)}{D(Z_c,\Theta)+\pi}$$ $$+ E(Z_c,\Theta)\cdot\pi^{F(Z_c,\Theta)} \quad \text{with}\ \pi = \pi(\tau)$$ | $(14)^d$ | Barile[7] | triple point to critical point. $A(Z_c,\Theta), B(Z_c,\Theta)\ldots$ are functions. |
| 12 | $$\rho_v = \frac{\mathcal{M}\, p_c\, \pi}{\mathcal{R}\, T_c\, \tau}\cdot\left\{1-(1-Z_c)\left[\frac{\Theta_1\tau^{200}+\Theta_2\tau^{\Theta_3-1}+}{(1-\Theta_1-\Theta_2)^2\,\tau^{\Theta_3}}\right]\right\}$$ | 3 | Kud et al.[8] | triple point to critical point |

[a]Abbreviations ρ: vapor density, T: absolute temperature, τ: reduced temperature, Z: compressibility factor for saturated vapor, Θ: vector of model parameter, $A(Z_c,\Theta), B(Z_c,\Theta)\ldots$ are functions, $\mathcal{R}$: universal gas constant, $\mathcal{M}$: molar mass, subscripts v: vapor, c: critical state, r: reduced.   [b]SRK = Soave Redlich Kwong. [c]PR=Peng Robinson, EoS: Equation of State. [d]Parameters already estimated.

Model 11 in Table 1 is Barile's correlation[7], which is based on modeling the compressibility factor on the saturated vapor line. The 14 estimated parameters are obtained for 23 compounds with different polarity values. The vapor density is calculated from the reduced pure component vapor pressure $\pi(\tau)$, the reduced temperature τ and the reduced compressibility factor Z_r for the saturated state. The only independent variables in this model are the reduced temperature and critical compressibility factor Z_c. This means that no further parameter estimation is necessary when using this model.

Model 12 is also based on modeling the compressibility factor at saturation (dew point lines). Concerning the parameters that have to be estimated, some of them have to be determined

arbitrarily. For example in model 12, one of them, is an exponent with a value equal to 200. This kind of determination depends on the chemical compound. Model 12 suffers from the fact that most of the correlations are based on polynomial expressions. This leads to the fact that the vapor densities are poorly estimated within the vicinity of the triple and critical points. This was one of thes reasons which motivate the development of a new semi-empirical equation for describing the vapor density of compounds. This correlation has the same advantages of the original one[8] (Model 12) and permits more accurate calculations in the vicinity of the triple and critical points.

The models (n° 8 to 10) are edited on the same base with the same equation. Different values of compressibility factors can be used in the models n° 8 to 10. They can depend on the equation of state selected (SRK or PR EoS). Details are described in Section 3.

Briefly summarized:

- Models 2 to 11 are compared in respect of performance.

- Models 2 to 10 have compound specific parameters that are estimated based on measurements.

- Model 11 no parameters have to be estimated.

Vapor density data at saturation are required to fit the model parameters which is now described.

2 DATABASE DESCRIPTION

Seventeen compounds of interest for industrial applications were selected, with polarity values for the dipole moment μ ranging from 0 to 9.7 x 10^{-30} Cm and are reported in Table 2. In addition, care was taken to ensure that the compounds varied considerably with regard to their chemical structures and chemical functionalities (e.g., acids, aromatics, alkanes, ketone, alcohols).

Pure component critical properties, vapor pressures and vapor densities were obtained from the literature[10]. Additional pure component vapor pressure and vapor density values published in Thermolit[39] and other references listed in Table 2 were also used. A simple thermodynamic consistency test was applied to use the best a priori data. The compressibility factor for vapor was calculated from the known pure component vapor pressure eq 1 and the vapor density data from the literature. The compressibility factor value was required to be between the critical compressibility factor and 1. Moreover, the calculated compressibility factors must decrease monotonically with increasing temperature. All vapor density data at saturation that did not fulfill these conditions were not considered for the data treatment. At this point we would like to mention that the test was more difficult for hydrogen chloride. The compressibility factor, calculated from the vapor density, showed a maximum function of temperature. Consequently, experimental data with very low vapor density values were excluded from the evaluation, because the compressibility factor is lower than the maximum. Vapor density data at very low pressures are experimentally very difficult to evaluate for compounds presenting very high tendency of adsorption on the wall of the measuring vessel. Probably, this is the reason why the vapor density data of hydrogen chloride have very different values, see the references in Table 2.

At this point the term "measurement data" used in this work has to be defined. In this paper, measurement data are: either really measured vapor density (experimental data) or vapor density calculated from an EoS. Such calculated measurement data could also be called "model based on directly measured vapor density data". However, the EoS used for this must also be parameterized directly on measured vapor density data. Fundamental equations of state are often used for this purpose. However, in most of cases EoS are parameterized using pure compound vapor pressure data by adjustment of alpha function parameters (e.g. according to Soave, Mathias and Copeman,

Twu, Coquelet et.al.[30]). Such calculated vapor densities by EoS were not used here. In effect in many case great deviations exist between calculated and measured vapor density[8]. It concerns also vapor densities calculated using equations of state that require acentric factor as an additional parameter in alpha function. If the parameterization of the EoS was based not only on vapor pressures but also include vapor densities, we would have more references concerning experimental vapor densities data as presented in Table 2.

Table 2 shows that the same number of saturated vapor density data points do not exist for each component. The data for each compound were used for parameterization using Models 2–10 and to test their predictive capabilities. Additional details and the predictive capability of each model are presented in Section 4.

Table 2. Property data of the pure components and sources of the experimental data used in this work.

n°	Compound	CAS registry number	T_c^a / K	p_c^b/ Pa x10⁻⁵	ρ^c / kg/m³	T_{Tp}^d / K	p_{Tp}^e / Pa	μ^g / Cm x10³⁰	References[g]	n° of vapor density data
1	Ammonia	7664-41-7	405.55	112.8	234.7	195.50	6060	5.0	10,11,12,13,39,42,44	41
2	Hydrogen chloride	7647-01-0	324.6	83.09	450	158.9000	13800	3.7	10,14,15,16,44	10
3	Sulfur dioxide	7446-09-5	430.8	78.834	521.77	197.69	1670	5.3	17,39,44	20
4	n-Hexane	110-54-3	507.5	30.12	232.9	177.83	1.188	0.0	16,40,44,47	8
5	Cyclohexane	110-82-7	553.64	40.75	273.0	279.7	5388	1.0	10,18,39,44	8
6	Ethylene	74-85-1	282.34	50.401	214.16	104	120	0.0	18,39,40	7
7	Chloro(difluoro)methane	75-45-6	369.3	49.89	496.277	115.76	0.3794	4.7	10,19,44	9
8	Benzene	71-43-2	562.05	48.95	305.0	278.68	4784	0.0	10,41,43	7
9	Toluene	108-88-3	591.75	41.26	292.0	178.18	0.03939	1.3	10,39,44	9
10	Chlorobenzene	108-90-7	632.35	45.2	365.4	227.95	8.446	5.3	20,21,22,23,24,25,39,44	29
11	Acetone	67-64-1	508.1	47	273	178.45	2.785	9.7	10,18, 39,44	9
12	Ethyl acetate	141-78-6	523.25	38.32	307.7	189.3	1.432	6.3	10,18,39,44,45	15
13	Diethyl ether	60-29-7	467	36.1	265	156.8	0.3954	4.3	10,18,39,44	8
14	Methanol	67-56-1	512.6	81.035	275.6	175.5	0.1863	5.7	10,18,39,44	7
15	Ethanol	64-17-5	513.9	61.48	276	159	7.190·10⁻⁴	5.7	10,44,46	9
16	Water	7732-18-5	647.10	220.64	322	273.16	611.7	6.2	10,20,43	45
17	Ethylene oxide	75-21-8	469	71.94	315	160.65	7.8	6.3	10,14,39,43,49	9

Concerning the pure component vapor density data at saturation, the assumed error was based on information available in the literature for vapor density measurements if no information is specified in the data source. Depending on the quality of the database, the relative uncertainty[31] $u_{r,\rho}$ or the so-called coefficient of variation[32,33] is between $\sim 0.01 - 0.05$. The greater uncertainty is assumed for small densities. These assumptions are based on the fact that apparatus is less accurate for measurement of vapor densities for a complete range of measurement (from the triple to critical points). The vapor densities have been omitted near the critical point due to the higher uncertainty of these values. Moreover, the new model will be tested for its suitability with cubic EoS. In effect, it is well known that cubic EoS show large deviations in the vicinity of the critical temperature. This is also the reason why we neglected these values. The uncertainty for the vapor density is calculated according to the equation[31] $u_\rho = u_{r,\rho} \cdot \rho_m$ where ρ_m is the measured vapor density.

A vapor pressure function with the corresponding parameter values is required for the new model 8 to 10. The pure component vapor pressure correlation must be available between the triple and critical points. The well-known Antoine equation[48] is not suitable because it can only calculate vapor pressure in a limited temperature range. The DIPPR 101 equation (DIPPR denotes the Design Institute for Physical Properties) was selected from an overview[39] to calculate the vapor pressure of the 17 substances. This equation was rearranged in such a way (see eq 1) that the critical temperature T_c and pressure p_c, as the known values, contribute to the calculation of the vapor pressure, p .

$$\pi = \frac{p}{p_c} = exp\left[\frac{\Theta_{p1}}{T_c}\left(1 - \frac{1}{\tau}\right) + \Theta_{p2} \cdot ln(\tau) + \Theta_{p3} \cdot T_c^{\Theta_{p4}}\left(\tau^{\Theta_{p4}} - 1\right)\right] = f(\tau, \Theta_p) \tag{1}$$

The parameters $(\Theta_{p1}$ to $\Theta_{p4})$ must be fitted. The data base for parameter estimation is listed in the references in Table 2. For more details concerning the method of parameterization the reader is invited to read Appendix 1.

Furthermore, we assumed that the measurement uncertainty for the vapor pressure and vapor density follow a Gaussian distribution under constant measurement conditions. This means that at a constant temperature, the measured value is a random variable with a Gaussian distribution. This trend is often observed in this kind of practice.

3 COMPUTATIONAL METHODS

3.1 The new model

As already described by Kud et al.[8], the derivation of the new model starts with the compressibility factor Z_v for the vapor state (eq 2):

$$Z_v = \frac{pv}{\mathcal{R}T}, \tag{2}$$

where v is the molar vapor volume and $\mathcal{R}$ is the universal gas constant. In eq 2 for the molar vapor volume, v, the relationship $v = \mathcal{M}/\rho_v$ with molar mass $\mathcal{M}$ and vapor density ρ_v is inserted with the reduced state variables τ and π. Then eq 3 is obtained for the vapor density

$$\rho_v = \frac{\mathcal{M}p_c}{\mathcal{R}T_c} \cdot \frac{\pi}{\tau Z_v}. \tag{3}$$

In eq 3 for the reduced vapor pressure π the empirical pure component vapor pressure described by eq 1 was used. Eq 4 is used as the new empirical model for the compressibility factor Z_v as a function $g(\tau, q, w, c)$ of the reduced temperature and the parameters q, w, c

$$Z_v = g(\tau, q, w, c) = Z_c + (1 - Z_c) \cdot \begin{cases} \left[1 - \left(\frac{\tau - c\tau_{Tp}}{1 - c\tau_{Tp}}\right)^q\right]^w & if \ N_z = 1, \tau_{Tp} \leq \tau \leq 1 \\ \sum_{i=1}^{N_z} a_i \left[1 - \left(\frac{\tau - c\tau_{Tp}}{1 - c\tau_{Tp}}\right)^{q_i}\right]^{w_i} & if \ N_z > 1, \tau_{Tp} \leq \tau \leq 1 \\ not \ defined & if \ \tau > 1 \end{cases} \quad (4)$$

where $w_i \in [0.1, 0.95] \subseteq \mathfrak{R}$, $q_i \in [1.1, 100] \subseteq \mathfrak{R}$, $a_i \in \mathfrak{R}$, $c \in [0.3, 0.998[\subseteq \mathfrak{R}$, $\tau_{Tp} = {T_{Tp}}/{T_c}$

is the reduced triple point temperature, T_{Tp} is the triple point temperature and N_z is the number

of terms in eq 4 and $\mathfrak{R}$ denotes real numbers. For the all the compounds investigated here, a single

term was sufficient, i.e., $N_z = 1$. The model with two terms presented by eq 4 was only used for

compound 2 describing the vapor density.

Parameters q, w describe the curvature of the compressibility factor function, and the

compressibility factor at the triple point is correctly estimated using parameter c. Furthermore, if

$c = 1$, the vapor phase at the triple point has the properties of an ideal gas. Thus, a parameter value

of c<1 should be applied to accurately describe the vapor density. This was achieved by limiting

the parameter estimation problem with regard to the parameters.

The vapor density (eq 5) can be summarized by combining eqs 1, 3 and 4.

$$\rho_v = \frac{\mathcal{M} p_c}{\mathcal{R} T_c} \cdot \frac{f(\tau, \Theta_p)}{\tau g(\tau, q, w, c)} \quad (5)$$

Eq 5 can be called semi-empirical because according to eq 2, it includes the state variable pressure

$f(\tau, \Theta_p)$ and compressibility factor $g(\tau, q, w, c)$ at the dew point line. Eq 5 requires a well-known

and reliable prediction of the pure component vapor pressure function $f(\tau, \Theta_p)$ which is also

parameterized considering the reduced triple point temperature τ_{Tp} (see eq A1 in Appendix 1). This

ensures more consistency with eq 4. In this way, the vapor pressure function, eq 1, and the

compressibility factor, eq 4, are linked. See Appendix 1 for more details. We denote eq 4 in conjunction with eq 2 as the "new model", or Model 8.

Some properties of the compressibility factor function (eq 4) and their comparisons with real behavior are discussed below. Eq 4 fulfils the necessary conditions for the compressibility factor at the dew point line:

$$\lim_{\tau \to \tau_{Tp}^{+}} g(\tau,q,w,c) \cong 1 \qquad\qquad \lim_{\tau \to \tau_{Tp}^{+}} \frac{d}{d\tau} g(\tau,q,w,c) \cong 0$$

$$\lim_{\tau \to 1^{-}} g(\tau,q,w,c) = Z_c \qquad\qquad \lim_{\tau \to 1^{-}} \frac{d}{d\tau} g(\tau,q,w,c) = -\infty$$

(6)

At the vicinity of the triple point temperature the one-sided limit from the right part of the function $g(\tau, q, w, c)$ goes to value close to 1. This means that the ideal gas state is reached. The corresponding 1st derivative to the reduced temperature goes to zero value. At the vicinity of the critical temperature the one-sided limit from the left of the function $g(\tau, q, w, c)$ goes to the expected critical compressibility factor Z_c and the corresponding 1st derivative to the reduced temperature goes to infinite value.

At this point of the model discussion, the parameters q, w, c of eq 4 are important. For this reason, we would like to present the estimated parameter values in this Section. These values of q, w, c are listed in Table 3. One can observe that for hydrogen chloride and n-hexane the values of c are lower than the other parameter's values. All other compounds show a value of close to 1. A very small value means that an important correction of the triple point temperature have to be considered. This may be an indication that the vapor density data should be checked more strictly. For hydrogen chloride this point was already mentioned in Section 2. For n-hexane, the database used for parameter fitting have also be checked.

Table 3. Parameterization results for the parameter q, w, c in eq 4 based on Monte-Carlo simulations[a].

n°	Compound	q	u(q)[b]	w	u(w)	c	u(c)
1	Ammonia	1.838695	0.021	0.532258	0.002	0.994401	0.007
2	Hydrogen chloride	2.529029	0.095	0.515845	0.012	0.300000	0.247
3	Sulfur dioxide	2.163975	0.015	0.537213	0.003	0.997772	0.002
4	n-Hexane	3.993721	0.065	0.537728	0.007	0.300000	0.385
5	Cyclohexane	1.907121	0.031	0.592908	0.002	0.988224	0.010
6	Ethylene	2.334183	0.021	0.543092	0.004	0.998000	0.845
7	Difluoro chloromethane	2.774430	0.027	0.548105	0.004	0.998000	1.098
8	Benzene	1.776046	0.012	0.498876	0.003	0.998000	0.074
9	Toluene	3.237345	0.023	0.606932	0.004	0.998000	1.324
10	Chlorobenzene	3.188483	0.029	0.597500	0.007	0.998000	0.182
11	Acetone	2.357968	0.010	0.544333	0.002	0.998000	0.848
12	Ethyl acetate	3.041948	0.047	0.553303	0.007	0.998000	0.789
13	Diethyl ether	2.140655	0.041	0.477395	0.006	0.998000	1.163
14	Methanol	2.850944	0.036	0.563624	0.004	0.995684	0.014
15	Ethanol	4.419733	0.120	0.529832	0.005	0.989826	0.044
16	Water	2.461485	0.008	0.551213	0.002	0.998000	0.026
17	Ethylene oxide	3.086286	0.064	0.592708	0.012	0.998000	2.641

[a] q, w, c are mean values. Exact simulations require more decimal places than the calculated uncertainty indicates. [b] u: uncertainty

This model can also be used for the parameterization of a cubic EoS, as described in ref[8]. Instead of using the specific critical compressibility factor, Z_c, for a compound, the critical compressibility factor of the cubic EoS is used. Appendix 2A show the parameters q, w, c obtained using the compressibility function according to eq 4 in the Γ SRK (Soave Redlich Kwong) and Γ PR (Peng Robinson) equations of state. The Γ function corrects the vapor pressure and vapor density (see Kud, Körkel, Maixner[8] for more details). Thus, the gaseous density and vapor density can be calculated as a function of T and p with an error less than 1 %.

3.2 Presentation of the evaluation criteria

In this Section, we will test the quality of the new model 8 for determining the vapor density. First, the test methods and the results of evaluation are described. Since there are 17 compounds with 8 different test methods, there exist a lot of intermediate results. The intermediate results for each chemical compound are discussed in the Appendix. Models 2 to 7 and 11 presented in Table 1 are compared with the new model 8.

The examination of the new model is based on statistical methods which are divided into three groups:

(A) the Goodness of Fit

(B) the Goodness of the parameter and

(C) the consistency tests.

In group A the Goodness of Fit of a model is examined using two test methods. The first test is based on the chi-square (χ^2) distribution. In the context of examination of different models with different numbers of parameters, the condition of application of AIC criterion of Akaike[26] is evaluated. This criterion represents to the goodness of fit. A necessary condition for this test is that the random error follows a standard distribution with a homogeneous (constant) variance, and therefore, the residual sum of squares is used as an objective function for the AIC criterion. In our case, the random error do not belong to the same standard distribution with a constant variance value. In this case, the sum of standardized squared residuals (we call it the sum of weighted squared residuals, SWS) with inhomogeneous variances was chosen. Consequently, the AIC criterion could not be used in this work.

The second test in group A deals with the predictive capability of the model. This checks the capability of a model to predict existing experimental data that have not been included in the set

of measurements for parameters estimation. Thus, a parameter estimate is calculated with only a part of the experimental data. The other experimental data, i.e. vapor density data are used for validation.

In group B, the quality of the estimated parameters is evaluated by testing the capability of the model (and its parameters) to predict vapor density data outside the range of temperature of experimental data used for the fitting. Concerning this test, the rank of the covariance matrix of the parameters and an analysis of its variance are realized. At this point it should be mentioned that a non-zero full-rank value is a prerequisite for an analysis of parameter variance. For this reason, the rank test is the only one for which no H_0 hypothesis with confidence region can be postulated.

In group C the thermodynamic consistency within the Domain of Definition (DoD) and outside the DoD will be checked. It is important to pass test B before evaluation of Test C. A consistency test must be performed; this is important for saturated vapor density interpolation within the DoD range and if extrapolation outside the DoD range. The predictive capabilities of the properties for the consistency test can be calculated only if the lower temperature limit of the database is above the triple point temperature and if the reduced temperature range for prediction is greater than 0.05. This is not possible for compounds 1, 3, and 16 (ammonia, sulfur dioxide and water). In this case a mark of zero point is awarded for every model. Table 4 shows the point allocation system.

Table 4. Evaluation criteria and point allocation for Models 2 to11.

Test method number and test statistic	H_0 hypothesis and scores

	accepted	refused
A Goodness of Fit		
1 Test of the χ^2 distribution	1	0
2 Predictive Capability	1	0
B Goodness of the Parameter		
3 Full-rank of the covariance matrix of the Parameters[a]	1	0
4 Parameter variance based on confidence region[a,b]	1	0
C Consistency Tests		
5 Within the Domain of Definition (DoD)		
5.1 Range test for Z. The compressibility factor Z for saturated vapor must fulfill the condition: $Z_c^c \leq Z < 1$.	1	0
5.2 1st derivative test for Z. The 1st temperature derivative must fulfill the condition: $dZ/dT < 0$.	1	0
6 Outside the DoD – The Predictive Capability		
6.1 Range test for Z. The compressibility factor Z for saturated vapor must fulfill the condition: $Z_c \leq Z < 1$.	1	0
6.2 1st derivative test for Z. The 1st temperature derivative must fulfill the condition: $dZ/dT < 0$.	1	0

[a]This test is not possible for the Barile-based model 11 in Table 1.

[b]If hypothesis H_0 based on confidence region is fulfilled for every parameter of the model.

[c]Compressibility factor at critical state.

For all test methods, without the rank test number 3, a H_0 hypothesis is formulated on the base on a distribution function or boundaries of physical variables. In the case of the rank test, there are

only two options. Either full rank exists or not. A confidence interval of 99 % is selected concerning for the statistical test quantity (for more information see Appendix 2B). One point is given for each individual test number and for each run from test 1 to test 6. If the test criteria are fulfilled, they have a positive score; otherwise, zero points will be awarded. For statistical test quantities we prefer to add the different score and not the calculated probabilities. The reason is explained in Appendix 2C.

To evaluate trustworthiness of the scores (see Table 4), a small part of the topic "Parameter Estimation" of the following Section must be explained. The details are, of course, discussed in following Section 3.3. Calculation of all statistical quantities presented in Table 4 are obtained from vapor density measurement data ρ_m with their corresponding uncertainties u_ρ (see Section 2). Thus, the uncertainty (standard deviation) of the experimental data influences the test evaluation. If only a single calculation of the points is performed with only one standard deviation u_ρ according to the equation $u_\rho = u_{r,\rho} \cdot \rho_m$, then only a single result is obtained for each test. Also, the standard deviation is a random variable with a χ^2 distribution, this means that if the measurement is repeated, another standard deviation is obtained. So, if the calculation of the H_0 hypothesis evaluations is carried out with this other standard deviation, then other results can be obtained. Since the given relative uncertainty $u_{r,\rho}$ do not contain any information about the statistical distribution, a distribution of the standard deviation u_ρ can be calculated using Monte Carlo (MC) simulations. For this purpose, a standard deviation is simulated for each experimental data in a measurement data set for each compound. The result is a data set of simulated standard deviations for each compound which is considered constant for a so defined MC run. The statistical quantities for test n° 1 to 6 are then calculated in a MC run with the experimental data and their corresponding simulated standard deviations. The assumed and rejected H_0 hypotheses are counted

for each MC run. One point is awarded for a compound and a model only if the H_0 hypothesis is fulfilled in at least 95 % of all the MC runs. The number of MC runs depends on the maximum accepted deviation of the total number of points for a model. This maximum deviation must be lower than 1 point for a model. Simulation experience reveals for this kind of calculations at least 500 MC runs are necessary to reach this desired uncertainty.

In the context of the described stochastic simulations it should be mentioned that for the prediction test n° 2 within a MC run a further random process has to be executed. In effect it is known that the order of the measurement data within a measurement data set for the parameter estimation is invariant on the result of the parameter estimation. However, in the case of test n° 2, the total measurement data set must be partitioned as previously described. Before the total data set of a compound must be randomized, since the measurement data from the literature are often sorted in ascending temperature order. For the multiple parameter estimates required in test n° 2, the currently simulated standard deviation data set is used. This data set has already been simulated for the MC run. More details are presented in Section 4.

3.3 Parameter estimation

First, the quality criterion for the model acceptance must be defined. The residual values are required for this calculation. The sum of the unweighted squared faults or the mean relative absolute percentage deviation (MRD, also known as AAD%) according to eq A14 in Appendix 14 is often used to evaluate the residual values. If the data exhibit a so-called double exponential distribution (Laplace's distribution) or display a large number of outliers, a robust estimate of the parameters based on the L_1 standard,[27] i.e., the MRD is recommended[27]. However, one difficulty

remains using MRD. There is no statistical test method to formulate an H_0 hypothesis based on a distribution function. If the measurements follow the standard distribution, which is often the case in this field of application, then parameterization based on the so-called L_2 standard (Euclidean norm), i.e., the likelihood estimator, should be used[28]. Maximizing the likelihood estimator is equivalent to minimizing the least square function[28]. So we used the least square function because the measured data followed a standard distribution. To this end, the sum of standardized squared residuals or we call it the sum of the weighted squares $SWS_{M,j}$ was calculated for each compound j and every model M using eq 7

$$SWS_{M,j} = \sum_{i=1}^{N_{m,j}} \left(\frac{\bar{\rho}_{m,ji} - \rho_{M,j}(\tau_i, \Theta_{M,j})}{\tilde{s}_{ji}} \right)^2, \tag{7}$$

where N_m denotes the number of measurement points. The function $\rho_{M,j}(\tau_i, \Theta_{M,j})$ is the vapor density calculated according to the model equation (Models 2 to 11 in Table 1) and compounds $j=1...17$. $\Theta_{M,j}$ represents the model parameters for compound j, $\bar{\rho}_{m,ji}$ is the mean of measured vapor density and $\tilde{s}_{ji}$ is the associated standard error of the mean $\bar{\rho}_{m,ji}$ at measurement point i. If the sum of the weighted squares follows a χ^2 distribution with $N_m - N_\Theta$ degree of freedom, the model is good. The proof of χ^2 distribution can be used to calculate a statistical probability to formulate a H_0 hypothesis for the Goodness of Fit. Note: In many textbooks of statistics, so-called "single event" measurements $\rho_{m,ji}$ are given in the objective function (eq 7). The variance s_{ji} belongs to these measured values. If, however, a mean value $\bar{\rho}_{m,ji}$ is given for the measurement, this includes the so-called standard error of mean $\tilde{s}_{ji} = {s_{ji}}/{\sqrt{n_i}}$ where n_i is the sample size. (see also textbook by Larsen and Marx[29] on page 321). Eq 7 is also a Maximum-Likelihood Estimator (MLE). The proof is described in Appendix 2E.

In the next step we have to solve the problem concerning the standard deviation s_{ji} or uncertainty of experimental data for the eq 7. This kind of standard deviation is a random variable which is unknown for most cases of literature data set. As described in the previous Section, the calculated standard deviation $s_{calc,ji} = u_{r,ji}\,\bar{\rho}_{m,ji}$ from the coefficient of variation $u_{r,ji}$ for the vapor density $\bar{\rho}_{m,ji}$ presented in Section 2 isn't a random variable because the coefficient of variation $u_{r,ji}$ is often a mean value from several measurement series. On the basis of Monte Carlo simulation a stochastic standard deviation s_{ji} can be simulated from the calculated standard deviation $s_{calc,ji}$. This is described in detail in Appendix 2D. Such MC simulations generates optimally χ^2 distributed variances which can be used. If the vapor densities given in the literature are trustworthy, the model description is insufficient if the H_0 hypothesis based on a χ^2 distribution test is rejected.

The parameter estimation problem [PE 1] is based on the likelihood probability with standardized residuals and the resulting optimization problem can generally be formulated for the models 2 to 10 as follows:

$$\min_{\Theta_M,\rho_M} \left(\bar{\rho}_m - \rho_M(\tau,\Theta_M)\right)^{\mathcal{T}} \Omega^{-1} \left(\bar{\rho}_m - \rho_M(\tau,\Theta_M)\right) \qquad \text{[PE 1]}$$

$$s.t.\ \rho_M(\tau_i,\Theta_M) > 0 \quad (i = 1,2\,...\,N_m, M = 2 - 10\,)$$

$$1.1 \leq \Theta_{M,1} \leq 100,\ 0.1 \leq \Theta_{M,2} \leq 0.95\ (M = 8 - 10)$$

$$0.3 \leq \Theta_{M,3} < 0.999\ (M = 8 - 10)$$

where $\Omega = diag\left(\tilde{s}_1^2, \tilde{s}_2^2\,...\,\tilde{s}_{N_m}^2\right) = diag\left(s_1^2/n_1,\ s_2^2/n_2\,...\,s_{N_m}^2/n_{N_m}\right)$ is the diagonal matrix of the simulated standard variances of mean and $\mathcal{T}$ means the transpose of a vector or a matrix. $\left(\bar{\rho}_m - \rho_M(\tau,\Theta_M)\right)$ is the vector of residuals. The formulated inequalities must also be fulfilled.

Note that for this kind of measurements, variance inhomogeneity exists because the measurement range of density extends over several orders of magnitude.

All calculations in this work were done using PTC Mathcad 15 software.

4 RESULTS AND DISCUSSION

The outline of this Section refers to Table 4.

A) Goodness of Fit

Test 1: Test of the χ^2 distribution

The parameters $\Theta_{M,j}$ were calculated for Models 2 to 10 and for each compound. For Model 11, only the temperature and critical compressibility factor, Z_c, were required. The defined test quantity is the probability $P_\alpha^{(\chi)}$ (eq 8), which is based on the χ^2 distribution.

$$P_\alpha^{(\chi)} = 1 - P(X < SWS, f) = P(X \geq SWS, f). \tag{8}$$

$P_\alpha^{(\chi)}$ is the probability of finding a random χ^2 variable, X, from the statistical population that is greater than or equal to the test quantile, SWS. X is a random χ^2 variable, and f is the degree of freedom, $f = N_m - N_\Theta$, where N_Θ is the number of parameters to be adjusted.

The formulation of the H_0 hypothesis based on eq 8 refers to the so-called single-sided test and serves for better understanding only. In the present case the statistic recommends a partitioning of the significance level into two areas of the χ^2 probability density function[29]. The two-sided test is the result and the associated H_0 hypothesis can be formulated as follows:

$$H_0: \frac{\alpha_0}{2} \leq P^{(\chi)}\left(\chi^2_{\frac{\alpha_0}{2},N_m-N_\Theta} \leq SWS \leq \chi^2_{1-\frac{\alpha_0}{2},N_m-N_\Theta}\right) \leq 1 - \frac{\alpha_0}{2}. \tag{H$_0$1}$$

The big advantage for this two-sided test is that the lower limit of probability $P_{\alpha/2}^{(X)}$ gives information about the so-called overfitting. If this limit is fallen below $\frac{\alpha_0}{2}$, then overfitting is present and the model is rejected. If the upper probability limit $P_{1-\alpha/2}^{(X)}$ is exceeded $1 - \frac{\alpha_0}{2}$, the model description is inadequate and the model is rejected, too. This way of testing is very strict but we get information about the true model suitability.

The individual results for each component with their interpretations can be found in the Appendix. The summarizing results regarding the models will be discussed in the main part. This applies generally in the following. The results concerning the χ^2 distribution (test n° 1) of fit for every compound and every model are listed in Table S3.1 in Appendix A3. Table 5 presents the results of test n° 1 of the rating for each model tested and evaluated.

Table 5. Model assessment for models defined in Table 1: Goodness of Fit.

		Model number									
n°	Test	2	3	4	5	6	7	8	9	10	11
1	Chi Square distribution based on [$H_0$1]	9	7	7	9	1	3	13	14	13	5
2	Predictive capability[a] based on [$H_0$2]	3	5	7	3	0	2	5	6	6	nd[b]
	Goodness of Fit (sum of scores)[c]	12	12	14	12	1	5	18	20	19	nd

[a]within the Domain of Definition. [b]not defined. [c]maximum sum of scores for every model: 34

As shown in Table 5 for test n° 1, Models 6 and 7 with the lowest scores were developed to represent both the vapor and liquid densities at saturation. The reference point is the critical point. Consequently, the models failed to represent the vapor density at low temperatures. It should be mentioned, that for some compounds negative values for vapor densities can be calculated between

two measurement points at lower reduced temperature, $\tau < 0.67$. The reason for this is that for some compounds a pole may occur. Models 6 and 7 were rejected for most compounds within their domain of definition. The Barile model (eq 11) is rejected for most compounds because it isn't sufficient rated according to our method of scoring.

Of the polynomial models only models 2 and 5 give satisfactory results. The other models 3 and 4 give unsatisfactory results. The reason for this not so good rating of the polynomial models is the so-called overfitting and therefore H_0 hypotheses [$H_0 1$] are rejected for many compounds. Characteristic for overfitting is a very small value for the objective function (SWS) which are listed in Table S3.2 in Appendix 3. Table 6 exemplifies the overfitting on compound 9 and the results for the other models are for comparison.

Table 6. Overfitting of the models 2 to 5 for the compound 9 (toluene) in test n° 1. Assessment based on SWS and the H_0 hypothesis [$H_0$1]. The other models are listed for comparison.

Model number	SWS	[$H_0$1][a]	Assessment
2	0.0383	0	overfitting
3	0.0018	0	overfitting
4	0.0030	0	overfitting
5	0.0003	0	overfitting
6	2393	0	inadequate
7	1699	0	inadequate
8	3.5	1	good fit
9	2.7	1	good fit
10	3.4	1	good fit
11	31.9	0	inadequate

[a]H_0 hypothesis: 1 model accepted, 0 rejected

The models 2 to 5 show very small values for SWS and dissemble a very good fit as also reported in Table S3.2 in Appendix 3. However, they do not belong to the required χ^2 distribution. That's why the H_0 hypothesis [$H_0$1] is rejected. Compound 9 in Table 6 is no exception. There are still compounds 5 and 11 which show the same results as 9. Summary for SWS:

"A very small SWS value is no guarantee for a very good fit assessment."

However, if the SWS is greater than the allowable quantile of the χ^2 distribution (it can be calculated by considering the degree of freedom, dof $= (N_m - N_\Theta)$), then it is the characteristic of the weakness of a model which is the case for models 6, 7 and 11. Therefore, the H_0 hypothesis is rejected for these models, too. If the sum of weighted squares is in the required range of the quantile for the χ^2 distribution, then the model is accepted. This is the case for models 8 to 10.

For the sake of completeness, reference is made here to statistical test variables used by many scientists who evaluate parameter estimates. These test quantities can't be considered in this work because no H_0 hypothesis can be formulated. These are the mean relative absolute percentage

deviation (MRDs) as described in eq A14-1 (Appendix 14) and the reduced root mean square error (RMSE) as described in eq A15-1 in Appendix 15 were also calculated for comparison. All values are calculated with the estimated parameter according to the definition [PE1] in Section 3.3 and are presented in Appendix 14 and 15. Concerning the reduced RMSE, Andrae[35] warns using this statistical quantity for model assessment. In this context, reference is made to a test variable for the calculation of the model bias, which is also frequently used (see table S16 in Appendix 16). For this test quantity no H_0 hypothesis can be formulated, too. But it should be noted that this test quantity is very sensitive to model bias.

Since MRD is a very frequently used evaluation quantity, a comparison with the hypothesis test [$H_0$1] should be made here using the example of overfitting presented in Table 6. Compounds 5, 9 and 11 with models 2 to 5 show overfitting status. Considering the hypothesis test [$H_0$1] (see Table S3.1 in Appendix 3) these compounds are strictly rejected as a statistical point of view. However, the MRDs show very small values and even suggest the best model evaluation (Table S14 in Appendix 14). Consequently, the overall result of the model evaluation in the last row of Table S14 is also influenced. Also, the models 2 to 5 show the best performance but it is totally contrary to the results presented in Table 5 for test n° 1. The results of MRD testing can be summarized as:

"A very small MRD value is no guarantee for a very good fit assessment."

The part of overfitting in relation to the overall rating in the last row of Table S14 is about 17 % (Out of 17 compounds show 3 compounds overfitting, that's 3/17~17 %) for almost each of the polynomial models 2 to 5, i.e. about 17 % (see Table S14 in Appendix 14). The failure not to recognize such a model weakness, e.g. overfitting, is one of the reasons why the Chi-square test

[H$_0$1] for assessment should be preferred. 17 % of test failure is non-negligible and this is probably the reason why the polynomial model performs better than the new model.

Test 2: Predictive capabilities

Next, we use the so-called k-fold cross validation test, a rigorous test for evaluating the models concerning the predictive capabilities. This test was applied within the domain of definition and was calculated during a Monte Carlo run with the current simulated measurement variances s_i^2 calculated. The measurement data are divided into two different classes. Class one of the data is used for parameter estimation. This estimated parameter vector is used to predict the remaining class to calculate the predicted sum of weighted squared residuals (PRE SWS). In this work, 10 to 14 % of the vapor density data points were to be predicted. This means that a so-called 10 fold cross validation test is aimed for the calculations in this work. Readers are referred to Appendix 4 for further details of the parameter estimation procedure. The PRE SWS is the quantile of the χ^2 distribution and is calculated according to

$$PRE\ SWS_{M,j} = \sum_{k=1}^{N_k} \left(\frac{\left(\bar{\rho}_{m,jk} - \rho_{M,j}(\tau_k, \Theta_{M,j}) \right)}{\tilde{s}_{jk}} \right)^2 , \tag{9}$$

where k is the index corresponds to the data of the jth compound used. The testing philosophy is the same as described previously. In this case, the following question must be answered: does the predicted test variable PRE SWS belong to the χ^2 distribution with $(N_k - N_\Theta)$ degrees of freedom? The hypothesis H$_0$ for the test of the χ^2 distribution is

$$H_0: \frac{\alpha_0}{2} \leq P_{PRE}^{(X)} \left(\chi_{\frac{\alpha_0}{2}, N_k - N_\Theta}^2 \leq PRE\ SWS \leq \chi_{1-\frac{\alpha_0}{2}, N_k - N_\Theta}^2 \right) \leq 1 - \frac{\alpha_0}{2} . \tag{[H$_0$2]}$$

If Hypothesis [H$_0$2] is valid, then the model has a high predictive capability. Thus, the prediction is a real forecast because it can be compared directly to the measured data. It should be noted that

this predictive capability test is a very rigorous test and it is very difficult to achieve a maximum score.

In the event that the degree of freedom is zero, the parameters are calculated and not estimated. This was the case for the compounds with seven experimental data points and six parameters to be estimated. Then, the seventh experimental data point is calculated. This was the case for compounds 6, 8, and 14 and Models 2 to 5. The Table S5 in Appendix 5 shows for test n° 2 the scores for every compound and every model. The results concerning the predictive capability for the models only are listed in Table 5 (test n° 2). As expected, models 6 and 7 have no predictive capability. All models do not show a satisfactory rating. The overall rating of the "goodness of fit" test from Table 5 shows that the new model performs the best of all models.

B) Goodness of the parameter

This Section examines the properties of the model parameters. A model consists of an equation and its parameters. The quality of the model has already been tested. Now the quality of the parameters should be checked. It indicates the variance of the estimated parameters. This is important for interpolations and extrapolations because the error of a calculated vapor density depends on the error of the parameters. First, the necessary condition is to be examined as to whether or not an error propagation can be expected at all. Then the errors of the parameters should be checked.

The covariance matrix calculation of the parameters is often based on the Jacobian matrix $J(\Theta) = -\frac{\partial \rho_M(\tau,\Theta)}{\partial \Theta^T}$. As the parameter estimation problem [PE 1] is constrained, higher terms have to be considered in the Taylor development of the covariance matrix. For this reason, the Hessian matrix is used in this work[27]. It also applies, of course, to the unlimited case in [PE 1].

$$H(\Theta) = \frac{\partial SWS(\Theta)}{\partial \Theta \, \partial \Theta^T} \qquad\qquad (10)$$

$$cov(\Theta) = 2H(\Theta)^{-1} \qquad\qquad (11)$$

However, in the case of MC simulations, the mean value and the variance of the parameters can be calculated directly from the estimated parameter values obtained from MC runs. This possibility was used for the new model proposed in this work. If the derivative-based parameter estimation [PE1] is close to one of the constraints related to a parameter, then there is no possibility of variation of this parameter. This means that no variance can be calculated for this parameter from the MC runs. In these cases a calculation can be made from the derivative function (e.g. Hessian matrix). This is a rough approximation, since only 2 terms of the Taylor development are considered for the Hessian matrix. It means that the given uncertainties of the estimated parameters are also rough approximations. The difficulties in the error propagation calculation are due to the formulation of the constrained parameter estimation problem [PE1] for parameter c in eq 4, which is correction at the triple point temperature. This is a weakness of the constrained parameter estimation if the optimization algorithm is too close to one of the constraints. The results of the estimated parameters and their uncertainties for model 8 are given in Table 3. For models 9 and 10, the results are listed in Appendix 2A. Also, for the evaluation of a single MC run the Hessian matrix is used.

Test 3: Full rank of the covariance matrix of the parameters

If measurements have no impact on a parameter, then the variance of this parameter is infinite and the rank of the covariance matrix is reduced. In this situation, at least one parameter cannot be determined by the measurements used for the data treatment. Rank reduction can also be due to

important experimental uncertainties during the measurement, in which case the measurement has no influence on the parameter. Another reason for rank reduction is highly correlated model parameters, which could arise due to a bad equation formulation could thus be an inherent problem in the model. In this case, one or more parameters must be set to a realistic constant. A deeper model analysis can identify the precise reason for rank reduction. In the event of a matrix rank reduction in any of the cases, the solution is not unique. It should be mentioned, as previously described, that the parameter covariance matrix is calculated from the estimated parameters for the current MC run. Table S6 in Appendix 6 shows the scores for every compound and model and Table 7 shows the results for the model assessment concerning the rank test (test n° 3). These Tables show the results of all MC runs.

Table 7. Model assessment for models defined in Table 1: Goodness of the parameters.

n°	Test	Model number									
		2	3	4	5	6	7	8	9	10	11
3	Rank of cov matrix[a]	15	8	17	17	17	17	17	16	17	nd
4	Parameter variance based on [$H_0$3][b]	2	0	0	1	17	17	15	11	10	nd
	Goodness of parameter[c]	17	8	17	18	34	34	32	27	27	nd

[a] of the parameters. [b] and based on χ2 test. [c]maximum sum of scores for every model:34

All models except Model 3 show very good results and thus the necessary prerequisite for calculating an error propagation on the parameters is fulfilled.

Test 4: Parameter variance based on confidence region

The aim was to test whether the calculated parameter variance $cov(\Theta)_{ii}$ for the parameter Θ_i was significantly different from a set value of variance $s_{0,i}^2$. This variance is a given reference known from measurements or experience[29]. The test quantity χ_t^2 was calculated according to

$$\chi_t^2 = \sum_{i=1}^{rank(cov\Theta)} \left(\frac{(rank-1)\cdot cov(\Theta)_{ii}}{s_{0,i}^2} \right), \tag{12}$$

where $rank - 1$ is the degree of freedom and the covariance matrix must have full-rank. The standard deviation (s_0) was calculated using the parameter value and its relative error ε_0 is defined to the following equation

$$s_{0,i} = \Theta_i\, \varepsilon_0. \tag{13}$$

The calculated standard deviation of the parameter is considerably much larger than the standard deviation of the measurement, which is very often observed in practice. Therefore, the relative error ε_0 should not be excessively small. We recommend accepting a relative error of 0.5 in practice which should be at least one magnitude greater than the relative error of the measurements

$$\varepsilon_0 = 0.5. \tag{14}$$

The test quantity χ_t^2 is χ^2 distributed. Therefore, the χ^2 test was used for the assessment. The probability $P_\Theta^{(\chi)}$ of all parameters of a model was calculated based on χ_t^2. The Hypothesis H0 for this test is as follows

$$H_0: P_\Theta^{(\chi)}\left(\chi_t^2 \le \chi_{1-\alpha_0, rank-1}^2\right) \le 1 - \alpha_0. \tag{[H$_0$3]}$$

The accepted significance level α_0 was also set to 0.01 as previously described. Now to the assessment. The results of the parameter variance in Table 7 test n° 4 show that the polynomial model 2 to 5 do not have satisfactory performance. The models 6 and 7 with two or three parameters show a very good rating. The new model 8 to 10 shows a satisfactory to good rating. For more details concerning the assessment of the compounds see Appendix 7.

The overall rating "goodness of parameter" indicates the necessary but not sufficient criteria for the ability to extrapolate the vapor density calculation into a non-measured temperature range. In the overall rating model 3 is not sufficient. Models 2, 4 and 5 are satisfactory. The new model shows a good to very good performance and the models 6 and 7 show the best performance. How can the result for models 6 and 7 be interpreted for a prediction of vapor density? The best way to do this prediction is to consider the Goodness of Fit rating. Models 6 and 7 show in this test (Table 5) no good performance at all. Consequently, prediction outside the domain of definition is probably incorrect, as we will also see in the following part related to consistency test.

C) Consistency Test

Test 5: Consistency within the Domain of Definition (DoD)

This test depends on the model that is to be validated. This test assesses whether the correlation is violating fundamental laws. In one case, the thermodynamic consistency of the compressibility factor is tested, and in the other case, its derivative properties are checked. Test n° 5 assesses whether the calculated saturated vapor density is consistent with the corresponding vapor pressure by calculating the compressibility factor of the saturated vapor phase. This model is taken because the temperature values of the pure compounds are known. A prerequisite for the test, as noted above, is that the estimated pure component vapor pressures must be accurately calculated. In this manner, the calculated saturated vapor densities from the model at the support points and the interpolation between the two support points can be checked for consistency with the derivative property.

The compressibility factor of a saturated vapor phase decreases monotonically with the temperature. Hence, we were able to formulate two test conditions for the compressibility factor $Z_{M,j}(\tau)$ of the vapor phase (eqs 15 and 16).

$$Z_c \leq Z_{M,j}(\tau) < 1 \tag{15}$$

$$\frac{d\,Z_{M,j}}{d\,\tau} < 0. \tag{16}$$

The test criterion was defined as the compressibility factor, $Z_{M,j}$, for Models 2 to 11 and compound j for the saturated vapor phase:

$$Z_{M,j} = \frac{\mathcal{M}_j\,p_{c,j}}{\mathcal{R}\,T_{c,j}} \cdot \frac{f(\tau,\Theta_{\mathbf{p,i}})}{\tau\,\rho_{M,j}}\,, \tag{17}$$

where $\rho_{M,j}$ is the vapor density model in Table 1. Of course, in the case of the new model, eq 4 is used directly as the test quantity. The Domain of Definition, D_j, was chosen to be equal to the temperature range in which the vapor density data are available: $D_j \in [\tau_{min,j}, \tau_{max,j}]$. For the model acceptance with regard to the compressibility factor, Hypothesis H_0 is

Test 5.1 $H_{0,j}$: $\quad Z_c \leq Z_{M,j}(\tau) < 1 \;\; if \;\; \tau_{min,j} \leq \tau \leq \tau_{max,j}.$ [H₀4]

For the model acceptance with regard to the 1st reduced temperature derivative of the compressibility factor, Hypothesis H_0 is

Test 5.2 $H_{0,j}$: $\quad \left(\dfrac{d\,Z_{M,j}(\tau)}{d\,\tau}\right) < 0 \;\; if \;\; \tau_{min,j} \leq \tau \leq \tau_{max,j}.$ [H₀5]

Table 8 shows the sum of the results for both H_0 hypotheses concerning the model assessment. The Tables S8 and S9 in the Appendix 8 and 9 show the individual results for each compound and Table S10 show the summarized scores.

Table 8. Model assessment for models defined in Table 1: Consistency test within the Domain of Definition (DoD)

n°	Test within the DoD	Model number									
		2	3	4	5	6	7	8	9	10	11
5.1	Range test for Za based on [H04]	13	16	16	16	6	7	17	17	17	17
5.2	1st derivative test for Z based on [H05]	9	13	13	10	4	1	17	17	17	17
	Consistency test (SoSb)	22	29	29	26	10	8	34	34	34	34

[a] compressibility factor at saturated vapor. [b] maximum Sum of Scores for every model:34

Both tests 5.1 and 5.2 are discussed together. As expected, Models 6 and 7 had difficulty describing the compressibility factor for the saturated vapor phase within the domain of temperature previously defined. The reason for this is that for some compounds a pole may occur (e.g. for the compound 14), which leads to singularity and thus to thermodynamic inconsistency and an insufficient rating. Model 2 is satisfactory. Models 3 to 5 show a good rating. The reason for the score deduction is the polynomial character of this model class, which causes oscillatory behavior. The results concerning the compressibility factor calculations will be shown as an example for the two models 2 and 5 in Figure S1, S2 in Appendix 8, 9. Figure S2 illustrate impressively the oscillating behavior of the polynomial model of model 5. It is of course ensured that the oscillating behavior is not due to the vapor pressure function $f\left(\tau, \Theta_{\mathbf{p}}\right)$ in the eq 5 because it is monotonically increasing. Both examples in Figure S1, S2 are accepted in the test 1 of the goodness of fit. Table 9 presents all the examples that have been accepted in test 1 (goodness of fit) for the H_0 hypothesis [$H_0$1] and that have been rejected in test 5 for thermodynamic consistency.

Table 9. Model assessment: Rejected thermodynamic consistency but accepted according to Goodness of Fit.

	Compound n°a						
Model n°b	1	3	6	7	14	15	17
2	oc	o	o	o	o	o	
3	o	o			o		
4	o	o			o		
5	o	o	o		o		
6							o
7							o

a compound name is listed in Table 2.

b model number is listed in Table 1.

c o means: [H01] accepted and [H04], [H05] refused.

This means that it is not only important to use statistical quantities for a model evaluation, but also fundamental physical quantities. In the case of testing kinetic models, it is the mass conservation law, and in the case of a thermodynamic model, it may be for example the compressibility factor. This test is highly sensitive to small inconsistencies and in all cases the new model shows thermodynamic consistency and therefore it shows a very good performance.

Test 6: Outside the Domain of Definition – The Predictive Capability

Here, we investigated whether the model is suitable for extrapolation beyond the domain of definition (DoD). The test range or definition range of prediction, DP_j, is between the triple point temperature and the smallest temperature value of the range of the definition $DP_j \in \left[\tau_{Tp,j} , \tau_{min,j} \right[$. The temperatures must be in the following range to ensure that the extrapolation range is still

physically meaningful: $\tau_{\min,j} - \tau_{Tp,j} \geq 0.05$. With these conditions, all compounds were suitable

with the exception of ammonia, sulfur dioxide and water. Hypothesis H_0 is

Test 6.1 $H_{0,j}$: $Z_c \leq Z_{M,j}(\tau) < 1$ if $\tau_{Tp,j} \leq \tau < \tau_{\min,j}$. [$H_0$6]

For model acceptance with regard to the 1st temperature derivative of the compressibility factor,

Hypothesis H_0 is

Test 6.2 $H_{0,j}$: $\left(\dfrac{d\, Z_{M,j}(\tau)}{d\, \tau} \right) < 0$ if $\tau_{Tp,j} \leq \tau < \tau_{\min,j}$. [$H_0$7]

Appendix 11, 12 show the individual results and its summarized scores are in Table S13. Table

10 shows the sum of the results for both H_0 hypotheses concerning test 6.1 and test 6.2.

Table 10. Model assessment for models defined in Table 1: Consistency test outside the Domain of Definition (DoD) – The predictive capability.

		Model number									
n°	Thermodynamic tests	2	3	4	5	6	7	8	9	10	11
6.1	Range test for Z^a based on [$H_0$6]	1	3	4	2	0	0	14	14	14	14
6.2	1st derivative test for Z based on [$H_0$7]	3	6	5	3	0	0	14	14	14	14
	Consistency test outside DoD^b	4	9	9	5	0	0	28	28	28	28

^{a} compressibility factor at saturated vapor.

^{b} maximum sum of scores for every model: 28

The results are discussed together for the tests 6.1 and 6.2. Models 6 and 7 show no predictive

capability. All polynomial models 2 to 5 show inadequate rating. One of the reasons for this is

the typical oscillating behavior of a polynomial model as shown in Figure S1 and S2 in

Appendix 8, 9. The new model 8 to 10 and Barile's model shows its best predictive capability in

thermodynamic consistency and a very good rating for all compounds in the case of an extrapolation outside the domain of definition.

Summary of the rating

   A summary of the rating based on Monte Carlo simulations as presented in Sections 2 and 3 was created by adding all scores of the six various test methods described in this Section and is shown in Table 11. The standard deviations of the total scores for Group 2 was calculated from Monte Carlo simulations and are listed in the last row of Table 11.

Table 11. Summary of the model rating based on Monte Carlo simulations.

Test Method and test number	Model number[a]									
	2	3	4	5	6	7	8	9	10	11
A Goodness of the Fit										
1 Test of χ^2 distribution	9	7	7	9	1	3	13	14	13	5
2 Predictve capability: Cross validation test	3	5	7	3	0	2	5	6	6	nd
B Goodness of the Parameter										
3 Full rank of the covariance matrix of the parameter	15	8	17	17	17	17	17	16	17	nd
4 Parameter variance based on the χ^2 Test	2	0	0	1	17	17	15	11	10	nd
C Consistency Test										
5 Consistency within the domain of definition	22	29	29	26	10	8	34	34	34	34
6 Predictive capability outside of the domain of definition	4	9	9	5	0	0	28	28	28	28
Group 1: Sum of scores from the test methods 1, 5, 6 [b]	35	45	45	40	11	11	75	76	75	67
Group 2: Sum of the scores of all the test methods.[c]	55	58	69	61	45	47	112	109	108	n.d.
Standard deviation for the scores in group 2	1.6	1.9	1.3	1.6	0.7	0.6	1.4	1.5	1.4	n.d.

[a]the corresponding model see table 1. [b]Maximum sum of scores for every model: 79. [c]Maximum sum of scores for every model: 130

Because many tests could not be calculated with the Barile model (Model 11), two groups for different model rankings were assigned as presented in Table 11.

Group 1: tests for minimization of the objective function (goodness of fit) and thermodynamic consistency (with Barile's model) exhibits the following model ranking: very good, good, satisfactory and unsatisfactory where the Symbol ">" denotes "better than"

Models 8 to 10 > Model 11 > Models 3, 4, and 5 > Models 2, 6 and 7

Group 2: like group 1 plus the tests for predictive capability and for "goodness of the parameter" (without Barile's model) exhibits the ranking: good, satisfactory and unsatisfactory

Models 8 to 10 > Models 4 > Models 2, 3, 5, 6 and 7

In both groups models 6 and 7 have an unsatisfactory ranking. The polynomial models 2 to 5 show a satisfactory to an unsatisfactory rating. The reason for this is explained in the Section before. Barile's model is good in group 1. If no vapor density data are available and only the test requirements for group 1 need be met, then this model can be recommended. However, due to the inadequate rating of the goodness of fit test, a larger deviation from the true value has to be accepted for the vapor density calculation.

If reliable vapor density measurements and a vapor pressure function for pure compounds are available then the new model 8 to 10 shows the best assessment in both groups. Therefore, this semi-empirical three-parameter equation for accurately predicting vapor densities is recommended.

The mean relative absolute percentage deviation (MRD, also known as AAD%) for a model was also calculated for comparison and analyzed for suitability as an assessment quantity. It has been shown that MRD exhibits, as expected, very small values for a very good fit. However, in about 17% of the cases, no overfitting is indicated for the evaluation of the polynomial models which have very small MRD values, in contrast to the strictly statistical evaluation. For this reason, this assessment was not considered. The H_0 Hypothesis test based on a Chi-square distribution reveals an overfitting without any problems in contrast to the MRD value. Therefore, for the model evaluation "Goodness of Fit", we recommend the formulation of an H_0 hypothesis based on the Chi-square distribution, if normally distributed measurement data are given. Of course, it is often the case for measurements in the field of research.

5 CONCLUSIONS

A new model has been developed to describe the vapor density of a pure compound. It is based on the modeling of the dew point line. It includes information about the vapor pressure and a function to describe the compressibility factor for the saturated vapor line. This function is parameterized with vapor density data and it needs only three parameters for the regression. This result is based on the stated relative uncertainty (coefficient of variation) of about 0.01 to 0.05 for the used data base. If a higher measurement precision is required, another term can be used in the new model. Then three additional parameters need to be estimated.

The quality of the new model has been tested in comparison to seven known models which represent the "state of the art". The validity of the model was tested and confirmed on nine different classes of compounds with very different polarity, e.g. aliphatic and cyclic hydrocarbons, ester, ether, ketone, alcohol and inorganic compounds. For the validation of the

model, only chemical compounds with trustworthy measurement error information were selected.

The examination of the model and the comparison with the "state of the art" is based on strictly statistical methods. For the test, random standard deviations of the vapor density data are simulated to avoid that the overall result of the evaluation does not depend only on a random event, Monte Carlo simulations were realized. The model evaluation was complemented by a thermodynamic consistency test based on the same statistical evaluation method. The scientist thus has the opportunity to select the equation according to the desired focus of the application. The statistical analysis shows the typical oscillating behavior of a polynomial model, e.g. the Wagner equation. The new model shows the best performance for the description of vapor density with respect to the goodness of fit and the thermodynamic consistency.

Another advantage of the model is that it can also be used in the Soave Redlich Kwong[36] and Peng Robinson[37] equations of state. This makes it possible to correct not only the vapor pressure with e.g. the α-function according to Soave but also the vapor density in a cubic equation of state. The parameters for this correction are provided for both cubic equations of state.

Notes

The authors declare no competing financial interest.

AUTHORS INFORMATION

Corresponding Author

E-mail: alexander.e.kud@yandex.com

Co-authors

E-mail: christophe.coquelet@mines-paristech.fr

E-mail: stefan.maixner@t-online.de

ORCID

Alexander Kud: 0000-0002-3637-1490

Christophe Coquelet: 0000-0001-6382-673X

Stefan Maixner: 0000-0002-3877-9246

REFERENCES

(1) Sengers, J.; Sengers, J. L. Thermodynamic Behavior of Fluids Near the Critical Point. Annu. Rev. Phys. Chem. **1986**, 37, 189-222.

(2) Coquelet, C.; Ramjugernath, D.; Madani, H.; Valtz, A.; Naidoo, P.; Meniai, A. H. Experimental Measurement of Vapor Pressures and Densities of Pure Hexafluoropropylene. J. Chem. Eng. Data **2010**, 55, 2093-2099.

(3) Guder, C. Eine Neue Zustandsgleichung für das Fluide Zustandsgebiet von Schwefelhexafluorid für Temperaturen von der Schmelzlinie bis 625 K und Drücke bis 150 MPa; Ruhr-Universität Bochum: Germany, 2006, 54-56.

(4) Funke, M.; Kleinrahm, R.; Wagner, W. Measurement and Correlation of the (P, ρ, T) Relation of Ethane Ii. Saturated-Liquid and Saturated-Vapour Densities and Vapour Pressures Along the Entire Coexistence Curve. Zh. Khim. Termodin. **2002**, 34, 2017-2039.

(5) Wagner, W.; Pruss, A. International Equations for the Saturation Properties of Ordinary Water Substance. Revised According to the International Temperature Scale of 1990. Addendum to J. Phys. Chem. Ref. Data 16, 893 (1987). J. Phys. Chem. Ref. Data **1993**, 22, 783-787.

(6) Hales, J. L.; Gundry, H. A. An Apparatus for Accurate Measurement of Liquid and Vapour Densities on the Saturation Line, up to the Critical Temperature. J. Phys. E: Sci. Instrum. **1983**, 16, 91-95.

(7) Barile, R. G.; Thodos, G. Saturated Vapor and Liquid Densities of Pure Substances. The Canadian Journal of Chemical Engineering **1965**, 43, 137-142.

(8) Kud, A.; Körkel, S.; Maixner, S. A Cubic Equation of State Based on Saturated Vapor Modeling and the Application of Model-Based Design of Experiments for Its Validation. Chem. Eng. Sci. **2010**, 65, 4194-4207.

(9) Walas, S. M. Phase Equilibrium in Chemical Engineering; Butterworths Tolley Limited: Stoneham, MA, 1985, 88.

(10) VDI-Wärmeatlas. Verein Deutscher Ingenieure e.V. Vdi-Gesellschaft Verfahrenstechnik und Chemieingenieurwesen, 13th ed.; Springer Verlag: Berlin, Heidelberg, 2013.

(11) Barthau, G.; Sohns, J. Thermische Eigenschaften von Ammoniak. Chem. Ing. Tech. **1974**, 46, 149.

(12) Cragoe, C. S.; McKelvy, E. C.; O'Connor, G. F. Scientific Papers of the Bureau of Standards. **1923**, 18, 707. Tab.5, 7.

(13) Ahrendts, J.; Baehr, H. D. Die Thermodynamischen Eigenschaften von Ammoniak; VDI-Verlag: Düsseldorf, **1979**; 9-19.

(14) Ahlberg, K. AGA Gas Handbook, by AGA AB, Lidigö, Sweden, 7th ed.; Almqvist & Wiksell International: Stockholm, Sweden, 1985.

(15) Schäfer, K; Lax, E. Landolt-Börnstein. Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik. XI, Bd. II: Eigenschaften der Materie in ihren Aggregatzuständen, 2. Teil: Gleichgewichte außer Schmelzgleichgewichten. Bandteil a: Gleichgewichte Dampf-Kondensat und osmotische Phänomene. 6. Aufl.; Springer-Verlag: Berlin-Göttingen-Heidelberg, **1960**. 189.

(16) Green, D. W.; Perry, R. H. Perry's Chemical Engineers' Handbook. 8th ed.; McGraw-Hill Education: New York, 2007.

(17) Walas, S. M. Phase Equilibrium in Chemical Engineering; Butterworths Tolley Limited: Stoneham, MA, 1985, 163.

(18) DDBST. Dortmund Data Bank Software & Separation Technology Gmbh; Center for Applied Thermodynamics, Marie-Curie-Str. 10, D-26129 Oldenburg, Germany. Online Service.

(19) Perry, R. H.; Green, D. W.; Maloney, J. O. Perry's Chemical Engineers' Handbook, 7th ed.; McGraw-Hill: New York, 1997, p2-273, Tab. 2-305.

(20) Zmaczynski, A.; Bonhoure, A. La Témperature D'ébullition de L'eau en Fonction de la Pression. J. Phys. Radium **1930**, 1, 285-291.

(21) Ramsay, W.; Young, S. LXVIII-A Method for Obtaining Constant Temperatures. J. Chem. Soc, Trans. **1885**, 47, 640-657.

(22) Deshpande, D. D.; Pandya, M. V. Thermodynamics of Binary Solutions. Part 2. - Vapour Pressures and Excess Free Energies of Aniline Solutions. Trans. Faraday. Soc. **1967**, 63, 2149-2157.

(23) Young, S. XLVIII-On the Vapour-Pressures and Specific Volumes of Similar Compounds of Elements in Relation to the Position of Those Elements in the Periodic Table. Part I. J. Chem. Soc., Trans. **1889**, 55, 486-521.

(24) Young, S. In The Vapour-Pressures, Specific Volumes, Heats of Vaporization, and Critical Constants of Thirty Pure Substances, Sci., Proc. Roy., Dublin Society, **1909-1910**; 12, 374-444.

(25) Agaev, T. S.; Kafarov, T. E.; Kerimov, A. M. Thermodynamic Properties of Halogenderivatives of Benzene. II. Density of Chlorobenzene. Zh. Fiz. Khim. **1976**, 50, 1901.

(26) Akaike H. A New Look at the Statistical Model Identification. IEEE Trans. Autom. Control. **1974**, 19, 716-723.

(27) Seber, G. A. F.; Wild, C. J. Nonlinear Regression. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons: New York, Chichester, Brisbane, Toronto, Singapore, 1988. Chapter 2.5.

(28) Sen, A.; Srivastava, M. S. Regression Analysis: Theory, Methods, and Applications; Springer-Verlag: New York, 1990. Chapt. 3.3.

(29) Larsen, R. J.; Marx, M. L. An Introduction to Mathematical Statistics and Its Applications, 2nd ed.; Prentice-Hall: New Jersey, 1986. Chapt. 10.

(30) Coquelet, C.; Chapoy, A.; Richon, D. Development of a New Alpha Function for the Peng-Robinson Equation of State: Comparative Study of Alpha Function Models for Pure Gases (Natural Gas Components) and Water-Gas Systems. Int. Journ. Thermophysics **2004**, 25, 133-158.

(31) Taylor, Barry N.; Kuyatt, Chris E. Guidelines for Evaluating an Expressing the Uncertainty of NIST Measurement Results. NIST Technical Note 1297, **1994** Edition.

(32) Hendricks, W. A.; Robey, K. W. The Sampling Distribution of the Coefficient of Variation. The Annals of Mathematical Statistics, **1936**, 7, 129-132.

(33) Sokal, Robert R.; Braumann, Carlos A. Significance Tests for Coefficients of Variation and Variability Profiles. Systematic Biology **1980**, 29, 60-66.

(34) Bevington, P. R.; Robinson, D. K. Data Reduction and Error Analysis for the Physical Sciences, 3rd ed.; Mc Graw Hill: New York, 2003. 195.

(35) Andrae, R., Schulze-Hartung, T., Melchior, P. Dos and don'ts of reduced chi-squared. Cornell University. Instrumentation and Methods for Astrophysics. 2010. https://arxiv.org/abs/1012.3754

(36) Soave, G. Equilibrium Constants from a Modified Redlich-Kwong Equation of State. Chem Eng. Sci. **1972**, 27, 1197-1203.

(37) Peng, D.-Y.; Robinson, D. B. A New Two-constant Equation of State. Ind. Eng. Chem. Fundam. **1976**, 15, 59-64.

(38) Draper, Norman R.; Smith, Harry. Applied Regression Analysis, 3rd ed.; John Wiley & Sons, Inc.: New York Chichester Weinheim Brisbane Singapore Toronto, 1998.

(39) National Institute of Standard and Technology (NIST). ThermoLit: NIST Literature Report based on NIST SOURCE Data Archive.

(40) VDI-Wärmeatlas. Verein Deutscher Ingenieure e.V. Vdi-Gesellschaft Verfahrenstechnik und Chemieingenieurwesen, 10th ed.; Springer Verlag: Berlin, Heidelberg, 2006.

(41) Goodwin, R. D. Benzene Thermophysical Properties from 279 to 900 K at Pressures to 1000 Bar. J. Phys. and Chem. Ref. Data. **1988**, 17, 1541–1636.

(42) Sandarusi, J. A.; Kidnay, A. J.; Yesavage, V. F. Compilation of Parameters for a Polar Fluid Soave-Redlich-Kwong Equation of State. Ind. Eng. Chem. Process Des. Dev. **1986**, 25, 957-963.

(43) Nelson, R. D.; Lide, D. R.; Maryott, A. A. Selected Values of Electric Dipole Moments for Molecules in the Gas Phase. National Standard Reference Data Series-National Bureau of Standards **1967**, 10, 9-34.

(44) Poling, B. E.; Prausnitz, J. M.; O'Conell, J. P. TheProperties of Gases and Liquids. 5th ed.; McGraw-Hill: New York, 2001. A20-A35.

(45) Landolt-Börnstein. Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik. II Band, 2.Teil, Bandteil a, 6. Aufl.; Springer-Verlag: Berlin, 1960. 204.

(46) Schroeder, J. A.; Penoncello, S. G.; Schroeder, J. S. A Fundamental Equation of State for Ethanol. J. Phys. and Chem. Ref. Data, **2014**, 43, 043102-1 - 043102-43.

(47) Bell, I. H.; Wronski, J.; Quoilin, S.; Lemort, L. Pure and Pseudo-pure Fluid Thermophysical Property Evaluation and the Open-Source Thermophysical Property Library CoolProp. Ind. Eng. Chem. Res., **2014**, 53, 62498-2508. (http://www.coolprop.org/index.html).

(48) Antoine, C. Tensions des vapeurs; nouvelle relation entre les tensions et les températures. Comptes Rendus des Séances de l'Académie des Sciences, **1888**, 107, 836–837.

(49) Grumbles, T.; Jackson, S.; Lenahan, R.; Reeser, D.; Stewart, J.; Szczepanski, D.; Townsend, D.; Viscomi, R.; Vogel, K.; Wagner, M.; Wincek, J.; Wise, G. Ethyleneoxide - Ethylene Oxide Product Stewardship Guidance Manual, 3rd ed.; Sasol North America, BASF Corporation, Bayer MaterialScience LLC, The Dow Chemical Company, BASF Corporation, Huntsman Petrochemical Corporation, Celanese Chemicals Ltd., Balchem Corporation/ARC, Lyondell Chemical Company, Old World Industries, Croda, Inc., Shell Chemical LP. **2007**. 3.

(50) Hogge, J. W.; Giles, N. F.; Knotts, T. A.; Rowley, R. L. ; Wilding, W. V. The Riedel vapor pressure correlation and multi-property optimization. Fluid Phase Equilib. **429**, 2016, 149-165.