Atmos. Chem. Phys., 11, 9431–9450, 2011
www.atmos-chem-phys.net/11/9431/2011/
doi:10.5194/acp-11-9431-2011

Atmospheric Chemistry and Physics

EVAPORATION: a new vapour pressure estimation methodfor organic molecules including non-additivity and intramolecular interactions

S. Compernolle, K. Ceulemans, and J.-F. Müller

Belgian Institute for Space-aeronomy, Ringlaan 3, 1180 Brussels, Belgium

Received: 21 February 2011 – Published in Atmos. Chem. Phys. Discuss.: 29 April 2011
Revised: 26 August 2011 – Accepted: 1 September 2011 – Published: 16 September 2011

Abstract. We present EVAPORATION (Estimation of VApour Pressure of ORganics, Accounting for Temperature, Intramolecular, and Non-additivity effects), a method to predict (subcooled) liquid pure compound vapour pressure p^0 of organic molecules that requires only molecular structure as input. The method is applicable to zero-, mono- and polyfunctional molecules. A simple formula to describe $\log_{10} p^0(T)$ is employed, that takes into account both a wide temperature dependence and the non-additivity of functional groups. In order to match the recent data on functionalised diacids an empirical modification to the method was introduced. Contributions due to carbon skeleton, functional groups, and intramolecular interaction between groups are included. Molecules typically originating from oxidation of biogenic molecules are within the scope of this method: aldehydes, ketones, alcohols, ethers, esters, nitrates, acids, peroxides, hydroperoxides, peroxy acyl nitrates and peracids. Therefore the method is especially suited to describe compounds forming secondary organic aerosol (SOA).

1 Introduction

The (subcooled) liquid pure compound vapour pressure p^0 of a molecule is an important property influencing its distribution between the gas and particulate phase. While the vapour pressure of hydrocarbons and monofunctional molecules follows simple relationships, that of polyfunctional molecules is more difficult to describe. This is partly because the vapour pressure of such molecules is typically lower and therefore the experimental error is larger, and partly because there are more complex interactions (inter- and intramolecular in the liquid, intramolecular in the gas phase) between the functional groups. The molecules comprising secondary organic aerosol (SOA), which is the focus of our research, are typically polyfunctional. These semi- and low-volatility molecules originate from the oxidation of volatile organic compound (VOC) and they are of such a large diversity that a full determination of all species is unrealistic, let alone that for each species a vapour pressure can be measured. Near-explicit volatile organic compound oxidation mechanisms, like the MCM (Master chemical mechanism Jenkin et al., 1997; Saunders et al., 2003), BOREAM (Biogenic compounds Oxidation and RElated Aerosol formation Model Capouet et al., 2008; Ceulemans et al., 2010), or the GECKO-A (Generator for Explicit Chemistry and Kinetics of Organics in the Atmosphere Aumont et al., 2005) aim to simulate the complex chemistry leading to oxygenated semi-volatile and low-volatile species. To simulate SOA formation, such a chemical mechanism can be coupled to a partitioning module, where it is typically assumed that these compounds partition to the condensed phase as a function of their vapour pressure. Frequently used is the equilibrium partitioning formalism proposed by Pankow (1994) where the organic aerosol is considered as a well-mixed liquid; although recent findings (Cappa and Wilson, 2011) suggest that also another mechanism is possible, where the aerosol is rapidly converted from an absorptive to a non-absorptive phase. Estimation methods are therefore desired, that can quickly but reliably calculate vapour pressure from basic molecular structure information (e.g. a SMILES (Simplified Molecular Input Line Entry Specification) notation).

For some vapour pressure estimation methods other molecular properties are required as input, such as the boiling point (Nannoolal et al., 2008; Moller et al., 2008; Myrdal and Yalkowsky, 1997). This is an advantage if this boiling point is experimentally known, but it can contribute to the overall error if it has to be estimated. Several estimation methods were developed primarily for the relatively volatile hydrocarbons and monofunctional compounds, rather than the

Correspondence to: S. Compernolle
(steven.compernolle@aeronomie.be)

low-volatility polyfunctional molecules. For example, for low-volatility compounds, the method of Joback and Reid (1987) overpredicts boiling points (Stein and Brown, 1994; Barley and McFiggans, 2010; Compernolle et al., 2010), and the method of Myrdal and Yalkowsky (1997) tends to overestimate vapour pressures (Barley and McFiggans, 2010) when provided with an experimental boiling point. Another frequently encountered limitation is that not all molecule types are covered by the method at hand. Therefore, we recently extended some estimation methods to cover e.g. hydroperoxides and peracids (Compernolle et al., 2010). Some methods assume additivity in $\ln p^0$ with respect to contributions from different functional groups (Capouet and Müller, 2006; Pankow and Asher, 2008), but this approximation breaks down especially for hydrogen bonding functional groups. The method of Moller et al. (2008); Moller (2010) includes a special term for alcohols and acids to address this issue. Both the methods of Nannoolal et al. (2008) and Moller et al. (2008); Moller (2010) include terms to describe group-group interactions. However, the number of groups needed to describe these interactions might become very large, with some parameters constrained by only a few molecules. Also the group interactions are described in a non-local way, i.e. the relative position of two functional groups does not matter, contrary to chemical intuition. Finally, recently new room temperature low vapour pressure data of polyfunctional compounds became available – especially diacids and polyfunctional diacids (Frosch et al., 2010; Booth et al., 2010, 2011) – and it turned out that the available methods do not predict this data well (Booth et al., 2010, 2011). For these reasons, a new estimation method addressing the above issues is desirable.

2 Data set

2.1 Data collection of vapour pressures and boiling points

The data used for the development of EVAPORATION is presented in Table 1 of the Supplement. Data can be present as (i) original experimental data, (ii) a pressure-temperature ($p^0(T)$) correlation – e.g. an Antoine equation or a Wagner equation-, (iii) a boiling point at atmospheric pressure or (iv) a boiling point at reduced pressure. Although original experimental vapour pressure data is preferable over a $p^0(T)$ correlation, the error due to the use of a $p^0(T)$ correlation within its appropriate temperature range is minor compared to other error sources. As collecting all individual points in a data file is time-consuming, this was not pursued in all cases, even when the original experimental data was available. When using a $p^0(T)$ correlation, we took points with a 10 K interval. For $p^0(T)$ correlations of secondary data sources, we took generally only the vapour pressures above 1 kPa (9.87×10^{-3} atm) into account. This follows the recommendations of the secondary data source Engineering Sciences Data Unit (ESDU). We adopted this procedure also for the other secondary references (such as Yaws, 1994; Poling et al., 2001, and the Korean Thermophysical Properties Databank, KDB), as we presumed that the lower end of the reported temperature range rather referred to the melting point, i.e. where a liquid vapour pressure is applicable but the given $p^0(T)$ correlation is not necessarily reliable.

Sublimation pressure data was converted to subcooled liquid vapour pressure data by taking into account the melting point temperature and enthalpy of fusion (see Sect. 2.2).

Boiling points at atmospheric or reduced pressure were assembled, mostly from Chemistry Webbook of the National Institute of Standards and Technology (NIST, Linstrom and Mallard) – with important contributions from the compilations of Weast and Grasselli (1989) and Aldrich (1990) – from Lide (2000) and Sanchez and Myers (2000). Hence most boiling points were from secondary sources.

The following groups of compounds can be distinguished: non-functionalized hydrocarbons, monofunctional compounds and polyfunctional compounds.

2.1.1 Non-functionalized hydrocarbons (alkanes and alkenes)

As their vapour pressures are generally considered to be well characterised, we made no attempt to retrieve the primary references for these compounds, and considered a single reference source per compound as being sufficient. The most important data sources were the books of Poling et al. (2001); Yaws (1994); Dykyj et al. (1999) and KDB. The data was always in the form of a pressure-temperature ($p^0(T)$) correlation. No aromatic compounds were considered, as they are beyond the scope of this work.

2.1.2 Monofunctional compounds

These include aldehydes, ketones, ethers, esters, peroxides, nitrates, peroxy acyl nitrates, alcohols, acids, hydroperoxides and peracids. For these compounds we tried also to collect the primary reference sources. As a rule, all primary reference sources for the same molecule were taken into account, and at most one additional secondary reference if (e.g. by chronology) it was clear that the secondary reference was not based on the primary reference sources. In addition to the data sources already mentioned for the non-functionalized hydrocarbons, important secondary data sources were ESDU, the compilations of Pankow and co-workers (Asher et al., 2002; Asher and Pankow, 2006; Pankow and Asher, 2008) and NIST. For secondary data sources the data was always in the form of a $p^0(T)$ correlation. ESDU is claimed to be of high quality and contains error estimations of the $p^0(T)$ correlation. Therefore it was preferred over other secondary references. If a $p^0(T)$ correlation or original experimental data set was available for a molecule, no boiling point or reduced pressure boiling

point was taken into account, as this point (most frequently from a secondary data source) would fall most frequently within the range of this correlation or data set. While for most monofunctional compound types data availability is satisfactory, for some, especially hydroperoxides, peracids and peroxy acyl nitrates, it is not.

2.1.3 Polyfunctional compounds

For bifunctional compounds the availability of vapour pressure data depends strongly on the molecule type. For diols and diacids the situation is best, with data for over 30 molecules and with often dozens of experimental data points per molecule, while for hydroxy nitrates and hydroxy acids data availability is very limited, with data for less than six molecules and often only in the form of a single data point. Also, not all group combinations are covered, e.g. we do not have vapour pressure data on carbonyl nitrates. Important secondary sources here are ESDU and NIST.

For compounds with more than two functional groups, availability is even a more severe problem, although specifically for functionalised diacids the situation improved in recent years thanks to efforts of the atmospheric community (e.g., Booth et al., 2010, 2011; Chattopadhyay and Ziemann, 2005; Soonsin et al., 2010; Cappa et al., 2007).

As opposed to monofunctional compounds, for polyfunctional compounds an available boiling point was taken into account even if a $p^0(T)$ correlation or original experimental data set was available, as the boiling point was generally above the range of this $p^0(T)$ correlation.

2.2 Conversion of sublimation pressure to subcooled liquid vapour pressure data

Sublimation pressures are converted to subcooled liquid vapour pressures by (e.g. Prausnitz et al., 1999)

$$\ln\left(\frac{p_l^0}{p_s^0}\right) = \frac{\Delta H_{fus}}{R}\left(\frac{1}{T}-\frac{1}{T_{fus}}\right) - \frac{\Delta C_{p,sl}}{R}\left(\frac{T_{fus}}{T}-1-\ln\left(\frac{T_{fus}}{T}\right)\right) \quad (1)$$

with p_l^0, p_s^0 the vapour pressures of the liquid and solid state respectively, R the ideal gas constant, ΔH_{fus} the enthalpy of fusion and $\Delta C_{p,sl}$ the difference between solid and liquid heat capacity. Note that in Eq. (1) T_{fus} is used instead of the (theoretically correct) triple point temperature, but this incurs little error. $\Delta C_{p,sl}$ is frequently not experimentally available and the estimation $\Delta C_{p,sl} \approx \Delta S_{fus} = \frac{\Delta H_{fus}}{T_{fus}}$ is used here. The conversion is especially relevant for the recent data on diacids and functionalised diacids (e.g., Booth et al., 2010; Chattopadhyay and Ziemann, 2005), where the temperature of measurement is far below T_{fus}. In case no experimental ΔH_{fus} and/or T_{fus} is available, it can be estimated by the simple method of Compernolle et al. (2011a). Note that the updated version of this method (Compernolle et al., 2011b) is not yet applied here.

Table 1. Illustration of the effect of branching on vapour pressure, by comparison of vapour pressure of some example branched hydrocarbons with their linear isomers. Branched hydrocarbons have a higher vapour pressure than their linear counterparts, except if the branching occurs on a double bond.

Branched molecule	Linear isomer	$\log_{10}\left(p^0_{br}/p^0_{lin}\right)$ (298 K)
Alkanes or alkenes with branching not on double bond		
2-Methyl propane	Butane	0.16
2-Methyl butane	Pentane	0.13
2,2-Dimethyl propane	Pentane	0.40
2-Methyl pentane	Hexane	0.15
3-Methyl pentane	Hexane	0.10
3,3-Dimethyl-1-butene	1-Hexene	0.36
4-Methyl-1-pentene	1-Hexene	0.16
Alkenes, branching on double bond		
2-Methyl-propene	1-Butene	0.008
2-Methyl-2-butene	2-Pentene	−0.02
2-Methyl-1-butene	1-Pentene	−0.02
2-Methyl-2-pentene	2-Hexene	0.007
2,3-Dimethyl-2-butene	2-Hexene	−0.09
2-Methyl-1-pentene	1-Hexene	0.02

2.3 Data weighting

Optimal parameters are obtained by multiple linear regression, such that

$$\sum_i w_i \left(\log_{10}\left(p^0_{est,i}\right) - \log_{10}\left(p^0_{exp,i}\right)\right)^2 \quad (2)$$

is minimised, with $p^0_{exp,i}$ the experimental vapour pressure data point i and $p^0_{est,i}$ the corresponding modeled vapour pressure. w_i is a weighting factor, introduced such that one molecule cannot dominate in Eq. (2), e.g. a molecule for which a large number of T, p^0 data points are available, as opposed to a molecule where only a single boiling point is available. We set arbitrarily that one molecule cannot weight more than $\eta = 3$ times more than another one. If $N_{data}(i)$ is the size of the data set of a certain molecule where data point i belongs to, then w_i is defined as

$$w_i = 1, \text{ if } N_{data}(i) \le \eta$$
$$w_i = \frac{\eta}{N_{data}(i)}, \text{ otherwise.} \quad (3)$$

Changing η between one (all compounds have equivalent weight, disregarding their data point number) and ∞ (all data points have equivalent weight) had only a minor effect on the final results.

2.4 Kovats retention indices from gas chromatography

From NIST, a large quantity of Kovats retention indices (RI) from gas chromatography (GC) are available. For nitrates and functionalised nitrates, collections are available from Fischer (1999); Kastler (1999). RI are calculated from retention times of the target molecule and of a set of linear alkane reference compounds. A simple and often used approach (e.g. Fischer et al., 1992) to calculate vapour pressure at 298 K from RI of the target molecule, is to use the correlation $\log_{10} p^0(298\,\mathrm{K})$ – RI of the reference compounds. An increase in RI by 100 then corresponds theoretically to a ~ 0.5 decrease in $\log_{10} p^0$ at 298 K. This approach presumes that target compound and reference compounds have the same affinity towards the column, which is not generally true. Furthermore, RI are measured mostly far above room temperature and – specifically for RI from temperature-programmed GC, as opposed to isothermal RI – not at one single temperature. Therefore we did not use RI for the parameter fitting of our p^0 estimation method. However, they are still used to draw qualitative conclusions.

2.5 Notes on specific molecule classes

2.5.1 Monofunctional carboxylic acids

Small carboxylic acids (∼1–5 carbon atoms) can undergo significant gas-phase dimerization. Acetic acid, for example, is known to be mostly in dimeric form at room temperature, but the effect weakens for larger molecules and higher temperatures. As the association effect is not incorporated in our model, the experimental data has to be corrected for this. The experimental vapour pressure is the sum of both monomeric (p^0_m) and dimeric (p^0_d) forms.

$$p^0 = p^0_\mathrm{m} + p^0_\mathrm{d} \tag{4}$$

$$K_\mathrm{assoc} = \frac{p^0_\mathrm{d}}{\left(p^0_\mathrm{m}\right)^2} \tag{5}$$

Therefore, the vapour pressure of the monomer p^0_m can be calculated from the experimental vapour pressure p^0 and the association constant K_assoc:

$$p^0_\mathrm{m} = \frac{-1+\sqrt{1+4p^0K_\mathrm{assoc}}}{2K_\mathrm{assoc}} \tag{6}$$

p^0_m is taken as observational data to fit the model. Association constants of small carboxylic acids are taken from Miyamoto et al. (1999).

2.5.2 Peroxy acyl nitrates

The only peroxy acyl nitrate for which a measured vapour pressure is available is peroxy acetyl nitrate (Bruckmann and Willner, 1983; Kacmarek et al., 1978). This hampers a cross-validation for this type of compounds. However, it is possible to estimate additional vapour pressures from Henry law's constants $H = p^0\gamma^\infty$ of peroxy acyl nitrates, under the assumption that the contribution to the infinite dilution activity coefficient γ^∞ of the peroxy acyl nitrate (PAN) group is the same as for peroxy acetyl nitrate itself. Assuming that $\ln\gamma$ of a peroxy acyl nitrate RPAN can be splitted into a contribution of the parent hydrocarbon RCH_3 and a PAN group contribution, one gets

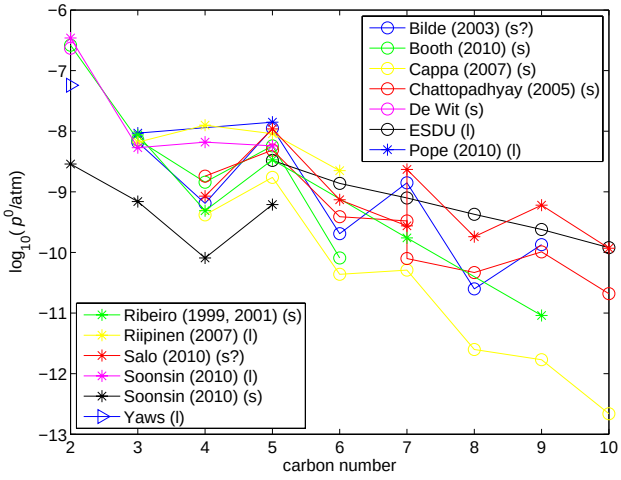

Fig. 1. $\log_{10}\left(\frac{p^0}{\mathrm{atm}}\right)$ vs. carbon number at 300 K for linear diacids, from different reference sources. (s) and (l) stand for solid and liquid, respectively. (s?) indicates that there is some doubt if all data was for solid particles.

$$\begin{aligned}\ln\gamma^\infty_\mathrm{PAN} &= \ln\gamma^\infty_\mathrm{RPAN} - \ln\gamma^\infty_\mathrm{RCH_3} \\ &= \ln\gamma^\infty_\mathrm{CH_3PAN} - \ln\gamma^\infty_\mathrm{C_2H_6}\end{aligned} \tag{7}$$

The vapour pressure of a general RPAN can be found from p^0, H data of hydrocarbons and of peroxy acetyl nitrate (CH_3PAN) and H data of the RPAN:

$$\begin{aligned}\ln p^0_\mathrm{RPAN} &= \ln\frac{H_\mathrm{RPAN}}{\gamma^\infty_\mathrm{RPAN}} \\ &= \ln H_\mathrm{RPAN} - \ln\gamma^\infty_\mathrm{RCH_3} - \ln\gamma^\infty_\mathrm{PAN} \\ &= \ln H_\mathrm{RPAN} - \ln\frac{H_\mathrm{RCH_3}}{p^0_\mathrm{RCH_3}} - \ln\frac{\gamma^\infty_\mathrm{CH_3PAN}}{\gamma^\infty_\mathrm{C_2H_6}} \\ &= \ln H_\mathrm{RPAN} - \ln\frac{H_\mathrm{RCH_3}}{p^0_\mathrm{RCH_3}} - \ln\frac{H_\mathrm{CH_3PAN}}{p^0_\mathrm{CH_3PAN}} \\ &\quad + \ln\frac{H_\mathrm{C_2H_6}}{p^0_\mathrm{C_2H_6}} \\ &= \ln\frac{H_\mathrm{RPAN}}{H_\mathrm{CH_3PAN}}\frac{H_\mathrm{C_2H_6}}{H_\mathrm{RCH_3}} + \ln\frac{p^0_\mathrm{RCH_3}}{p^0_\mathrm{C_2H_6}} + \ln p^0_\mathrm{CH_3PAN}\end{aligned} \tag{8}$$

Data of Henry's law constants was taken from Kames and Schurath (1995); Sander (1999).

2.5.3 Diacids and functionalised diacids

Recently, ambient temperature vapour pressure data of several research groups on diacids and functionalised diacids became available: Booth et al. (2010, 2011); Pope et al. (2010); Chattopadhyay and Ziemann (2005); Bilde et al. (2003); Monster et al. (2004); Frosch et al. (2010); Soonsin et al. (2010). This data is critical for the development of our vapour pressure method, primarily intended for polyfunctional molecules that are present in SOA. However, there can be orders of magnitude difference between measurements by different groups for the same compound, way above the reported experimental errors (typically 30–50 %). Figure 1 shows the vapour pressure vs. carbon number at 300 K for linear diacids calculated from $p^0(T)$ correlations of different reference sources, up to 10 carbon atoms. From 11 carbon atoms on, a departure of the expected vapour pressure or vaporisation enthalpy is observed, probably due to gas-phase cyclization (Ribeiro da Silva et al., 1999; Roux et al., 2005), and therefore this data is not included. Both liquid and solid data sets are present. Note that the shown points of ESDU, Yaws (1994) are obtained by bold extrapolation of $p^0(T)$ correlations from the appropriate temperature range. To a lesser extent this also applies to the data of Ribeiro da Silva et al. (1999, 2001). The data sets of subcooled liquid vapour pressures that are not extrapolations (Soonsin et al., 2010; Pope et al., 2010; Riipinen et al., 2007) agree relatively well with one another. Such data is most relevant to the parameterization of our method, as it is intended to predict liquid vapour pressures. Unfortunately, at room temperature liquid data is available only up to 6 carbon atoms, and no data is available for nonlinear or functionalised diacids.

The data for solids on the other hand shows severe disagreement, with the most extreme example being three orders of magnitude different for sebacic acid (ten carbon atoms) between the data of Salo et al. (2010) and Cappa et al. (2007). It has been speculated that this might be due to the experimental technique employed (Cappa et al., 2007; Pope et al., 2010) or to the physical nature of the diacids (Zardini et al., 2006; Soonsin et al., 2010; Salo et al., 2010) (presence of defects, partially or totally amorphous/liquid behaviour). Soonsin et al. (2010) also present vapour pressures of saturated solutions that should in theory equal the sublimation pressure of the corresponding crystalline solid particle, but without the complications encountered for solid particles. Inclusion of all available data in our model would lead to large uncertainties in the fitting parameters. Rather, we did a selection, although we are fully aware that the debate – which vapour pressure data set of diacids is the most reliable? – is not settled. For linear chains, we selected the liquid data sets (Soonsin et al., 2010; Pope et al., 2010; Riipinen et al., 2007) because of their mutual consistency. Second, the saturated solution data of Soonsin et al. (2010) was selected for malonic, succinic and glutaric acid, while we chose their solid particle data for oxalic acid. Soonsin et al. (2010) cite two reasons why their data on saturated solutions is more reliable than their data on solid particles themselves. First, the possible non-sphericity of solid particles bears uncertainty on measured vapour pressure, and second, the non-constant evaporation rates, probably due to liquid inclusions, complicates their measurements. For oxalic acid the solid particle data was chosen, as the saturated solution data was for the dihydrate rather than anhydrous oxalic acid. Another reason to choose the data of Soonsin et al. (2010) was its consistency with the corresponding liquid vapour pressure data and the fusion enthalpy. Finally, the sublimation pressure data of Cappa et al. (2007) was chosen as it is the most consistent with that of Soonsin et al. (2010) and extends to 10 carbon atoms.

For the nonlinear and functionalised diacids, no data from these references is available. We took therefore data from Monster et al. (2004); Booth et al. (2010, 2011); Ribeiro da Silva et al. (2000, 2001); Bilde and Pandis (2001); Chattopadhyay and Ziemann (2005); Frosch et al. (2010). The sublimation pressure data of the group of Bilde and coworkers (Bilde and Pandis, 2001; Monster et al., 2004) is relatively high, and we assume that they actually correspond to liquid vapour pressure, as it has been suggested before for the odd-numbered linear chain diacids (Zardini et al., 2006; Soonsin et al., 2010). High temperature (above the melting point) liquid vapour pressure data for diacids and a few functionalised diacids is taken from ESDU and Yaws (1994).

MacLeod et al. (2007) derived a linear relationship between ΔH_v and $\log_{10} p^0$ for non-hydrogen bonding compounds starting from Trouton's rule. Epstein et al. (2010) established a more general empirical linear relationship including also hydrogen-bonding compounds. It is informing to investigate whether the data on diacids and functionalised diacids obey this relationship. Figure 2 shows that while such a linear correlation is indeed observed for various compounds (alkanes, aldehydes, esters, alcohols, diols, hydroperoxides, peracids, peroxy acetyl nitrate and water were taken here), this is in general not the case for the diacids and functionalised diacids. The data of ESDU on diacids, and of Yaws (1994) on functionalised diacids does obey the correlation, notwithstanding the fact that the data points are bold extrapolations from the appropriate temperature range. Also the data of Cappa et al. (2007) obeys the correlation satisfactorily, and this is an additional argument why we chose their data as being representative for linear diacids. Many of the other data points, especially those of Booth et al. (2010); Monster et al. (2004); Bilde et al. (2003) are far from the correlation. This is in itself no proof that these data points are incorrect; for hydrogen-bonding compounds, the ΔH_v vs. $\log_{10} p^0$ relationship is empirical after all. But it does clearly show that the measured vapour pressure behaviour of these compounds strongly deviates from the expected pattern.

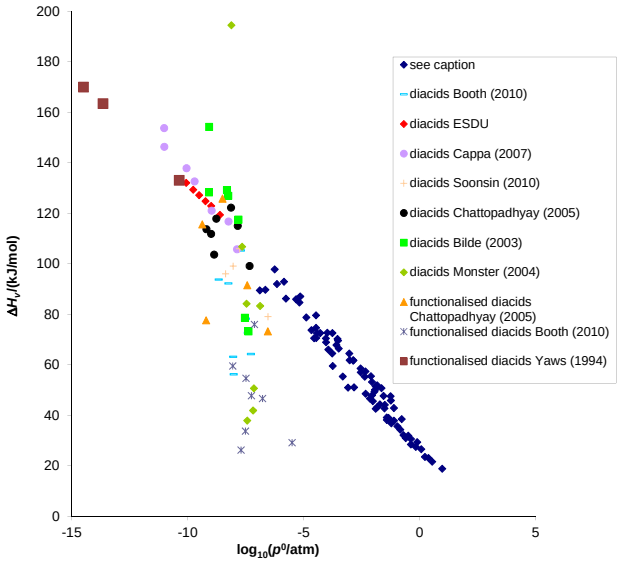

Fig. 2. ΔH_v vs. $\log_{10} p^0$ at 298 K for various compounds. The blue points serve as reference and include alkanes, aldehydes, esters, alcohols, diols, triols, hydroperoxides, peracids, peroxy acetyl nitrate and water. The other points are for diacids or functionalised diacids from various references, converted to subcooled liquid state, assuming $\Delta C_{p,ls} = \Delta S_{\text{fus}}$, if necessary. For a few compounds from Chattopadhyay and Ziemann (2005); Monster et al. (2004), ΔH_{fus} is estimated by the method of Compernolle et al. (2011a).

3 Statistical evaluators

Before describing the method framework of EVAPORATION and the procedure to fit its parameters, we will describe here the statistical evaluators that will be used to report the performance of the EVAPORATION. They include the model bias or mean deviation (MD), the mean absolute deviation (MAD), indicating the ability of the model to fit the data, and the predicted MD and MAD, indicating the predictivity of the model. MD and MAD will also be used to report the performance of other vapour pressure estimation methods on the molecules in our database.

$$\text{MD} = \frac{1}{N}\sum_{i=1}^{N} \log_{10} p^0_{\text{est},i} - \log_{10} p^0_{\text{exp},i} \tag{9}$$

$$\text{MAD} = \frac{1}{N}\sum_{i=1}^{N} \left| \log_{10} p^0_{\text{est},i} - \log_{10} p^0_{\text{exp},i} \right| \tag{10}$$

$$\text{pred.MD} = \frac{1}{N}\sum_{i=1}^{N} \log_{10} p^0_{\text{pred},i} - \log_{10} p^0_{\text{exp},i} \tag{11}$$

$$\text{pred.MAD} = \frac{1}{N}\sum_{i=1}^{N} \left| \log_{10} p^0_{\text{pred},i} - \log_{10} p^0_{\text{exp},i} \right| \tag{12}$$

with p^0_{est} obtained by fitting the model to all available data points i. Note that one molecule corresponds in general to several data points, from several reference sources and/or for several temperatures. $p^0_{\text{pred},i}$ is obtained by fitting the model to all data points, except those of the molecule which i belongs to. Note that to calculate these evaluators only a multiple linear regression (MLR) was performed; the few non-linear parameters (κ for the method optimised for zero- and monofunctional compounds, see Sect. 4.2, $\kappa, r, N_{\text{eff}}$ for the method optimized for all compounds, see Sect. 4.3) were kept fixed. The evaluators pred. MD and pred. MAD provide in terms of molecules a leave-one-out cross-validation (for each item to be estimated, its experimental value is left out of the fitting set, while all other values remain in the fitting set), but in terms of data points this is a leave-many-out procedure (as leave-one-out, but now for groups of items), as one molecule corresponds in general to several data points. Performing a separate MLR for each left-out molecule would be very inefficient: take a data set of 500 molecules, assume for simplicity that each molecule corresponds to 20 data points, and that 40 parameters are to be optimised, this would amount to solving 500 linear systems of size (10 000-20) × 40. Applying the work of Besalu (2001) on the leave-many-out method, the problem can be reduced to solving 500 linear systems of size 20 × 20. Specifically, Eqs. (6) and (7) of Besalu (2001) were used to calculate the $p^0_{\text{pred},i}$. Although Besalu (2001) divided the data set in portions of equal size for sequential prediction, we found that it was not necessary to do so.

4 Method outline

We first describe the temperature dependence of the method. Next, a method applicable to zero- and monofunctional compounds is described. Up to this level, the formulation follows that of a simple group contribution method. Then the method is extended to polyfunctional compounds, and it is described how non-additivity of functional groups is taken into account.

4.1 Temperature dependence

To describe the temperature dependence of the vapour pressure, the following simple empirical formula is proposed:

$$\log_{10} \frac{p^0}{\text{atm}} = A + \frac{B}{T^\kappa} \tag{13}$$

Basically the same formulation was presented by Korsten (2000), who adopted $\kappa = 1.3$, to describe the vapour pressure of hydrocarbons with or without hetero-atoms in a wide temperature range. Note that setting $\kappa = 1$ returns the basic Clausius-Clapeyron equation under assumption of a temperature-independent enthalpy of vaporisation -also known as the August equation-, valid only in a small temperature interval. A more precise description of the temperature evolution could probably be reached by introducing a larger number of group-specific coefficients, as in SIMPOL (Simplified p^0_L prediction method, Pankow and Asher, 2008), but

Eq. (13) was chosen for its simplicity and to avoid the possibility of overfitting.

The term A is directly related to the entropy of boiling at 1 atm total pressure $\Delta S_\mathrm{b} \equiv \Delta S_v(T_\mathrm{b})$, as from Eq. (13) it follows

$$T_\mathrm{b} = \left(-\frac{B}{A}\right)^{1/\kappa} \tag{14}$$

and, under the assumption of an ideal gas,

$$\frac{d\log_{10} p^0}{d(\frac{1}{T})} = -\frac{\Delta H_v}{\ln(10)\,R} = \frac{\kappa B}{T^{\kappa-1}} \tag{15}$$

Hence the enthalpy of vaporisation ΔH_v and of boiling ΔH_b is given by:

$$\Delta H_v = -\frac{\kappa \ln(10)\,R}{T^{\kappa-1}} B \tag{16}$$

$$\Delta H_\mathrm{b} \equiv \Delta H_v(T_\mathrm{b}) = A\kappa \ln(10)\,R T_\mathrm{b} \tag{17}$$

Combining Eq. (17) with the relation $\Delta H_\mathrm{b} = \Delta S_\mathrm{b} T_\mathrm{b}$ results in

$$A = \frac{\Delta S_\mathrm{b}}{\kappa \ln(10)\,R} \tag{18}$$

4.2 Method for zero- and monofunctional compounds

The most basic group-contribution approach describes $\log_{10} p^0$ as a sum of group contributions (Capouet and Müller, 2006; Pankow and Asher, 2008). This model is adequate for zero- and monofunctional compounds. A and B are then both divided into a sum of group contributions:

$$A = \sum_k c_k a_k \tag{19}$$

$$B = \sum_k c_k b_k \tag{20}$$

where a_k, b_k can be both first-order group contributions or second-order corrections on these group contributions. c_k are the values of a set of molecular descriptors. These descriptors are accountable molecular properties, obtained from molecular structure information. An important example is the number of times a certain functional group is present in a molecule. The first-order groups describe the molecule as a set of fragments (carbon atoms and functional groups), while the second-order groups take the environment of functional groups into account. The parameters a_k, b_k then connect the descriptor values to the observable estimate (here $\log_{10} p^0_\mathrm{est}$). Combining Eqs. (2), (19) and (20) results in

$$\sum_i w_i \left(\log_{10}\left(p^0_{\mathrm{est},i}\right) - \log_{10}\left(p^0_{\mathrm{exp},i}\right)\right)^2 = \sum_i w_i \left(\sum_k c_k a_k + \frac{\sum_k c_k b_k}{T^\kappa} - \log_{10}(p^0_{\mathrm{exp},i})\right)^2 \tag{21}$$

which is the function to be minimised. The problem is linear in the parameters a_k, b_k and thus can be solved by MLR at fixed κ. We report also the total group contribution g_k at 298 K, defined as

$$g_k = a_k + \frac{b_k}{(298\mathrm{K})^\kappa} \tag{22}$$

and its standard deviation

$$\sigma_k = \sqrt{\mathrm{covar}(a_k) + \frac{\mathrm{covar}(b_k)}{(298\mathrm{K})^{2\kappa}}} \tag{23}$$

with $\mathrm{covar}(a_k), \mathrm{covar}(b_k)$ the corresponding diagonal elements of the covariance matrix. To test whether the descriptor k is statistically significant, a student's t-test is performed: it was checked if

$$p\text{-value} = 1 - \int_{-u}^{u} f(t, \mathrm{df})\mathrm{d}t \tag{24}$$

with $u = g_k/\sigma_k$, $f(t, \mathrm{df})$ the student's t probability density distribution, df the degrees of freedom, and the p-value the probability that the null hypothesis is true, i.e. that g_k is not statistically different from zero. A high p-value (above the significance level) indicates that the null hypothesis cannot be rejected, and hence the descriptor was not retained. A significance level of 0.05 was taken.

To calculate a p-value from a student's t probability density distribution the degrees of freedom (df) has to be specified. The degrees of freedom are "the number of independent units of information in a sample relevant to the estimation of a parameter" (Everitt, 2010). Our approach is different from that of e.g. Raventos-Duran et al. (2010), where, df = #species − #parameters (or more generally, #observables − #parameters). As the number of species is much higher than the number of parameters, the distribution would then essentially become a normal probability density distribution, with a minimal width. In our opinion, this approach is too optimistic, probably only true when all observables are important to constrain all parameters. Taking as example the peroxy acyl nitrates, only a limited amount of information, namely data on 5 molecules, is available to constrain the parameter for the peroxy acyl nitrate group, the other data being irrelevant for this purpose. Instead, we define degrees of freedom as

$$\mathrm{df} = \#(\text{species where descriptor occurs}) - 1 \tag{25}$$

Hence df, as we define it here, is specific for each descriptor.

4.2.1 Size and topology of the molecule, evaluating hydrocarbons only

Apart from a constant term ($c_1 = 1$), two descriptors are used to describe hydrocarbons. As a descriptor related to the size of the molecule, the number of carbon atoms are counted; for functionalised molecules also the number of in-chain oxygen

atoms is counted. In-chain oxygen atoms are oxygen atoms that cannot be removed without breaking the carbon skeleton and occur in ethers (C**O**C), esters (C(=O)**O**C) and peroxides (C**OO**C). As a descriptor for the topology of the molecule, the topological index t is defined as

$$t = \text{branching number} - \text{ring number} \quad (26)$$

where the branching number is defined by taking at each carbon the number of *single* carbon-carbon bonds exceeding 2. The notion of single bonds is important as we found that branching at double bonds has no impact on the vapour pressure (Table 1).

As ring number and branching number have an impact on $\log_{10} p^0$ that is similar in magnitude but opposite in sign, we lumped them into the single descriptor t. With the few descriptors given above, all non-functionalised hydrocarbons in our database (130 molecules) can be described. Performing the regression for several κ an optimal value (smallest STD) for $\kappa = 1.5$ was found, somewhat higher than the value proposed by Korsten (2000). The method performs well for hydrocarbons, with an MAD of 0.057 and a pred. MAD of 0.060.

4.2.2 Including functional groups and local structure effects, evaluating also monofunctionals

Adding the monofunctional compounds to our fitting set results in a total of 579 species. $\kappa = 1.5$ was still the optimal value. An overview of the descriptors, together with their optimal parameters a_k, b_k for hydrocarbons and monofunctional compounds is given in Table 2. Also given in Table 2 is the total group contribution at 298 K g_k, and the combined standard deviation.

Parameters are introduced for the functional groups nitrate, carbonyl (including both aldehydes and ketones since their vapour pressures are very similar), ester, peroxy acyl nitrate, hydroxyl, acid, hydroperoxide and peracid. Note that with the functional group "carbonyl" we designate aldehydic and ketone groups, but not e.g. esters or carboxylic acids. Ethers and peroxides have no separate functional group contributions, as they are already accounted for by descriptor $k = 2$. Note that the hydrogen bonding groups (hydroxyl, acid, hydroperoxide, peracid) have about the same high a value of ~ 1. In other words, they give a similar contribution to the entropy of boiling. The high value is due to the higher ordering in the liquid phase compared to non-hydrogen bonding liquids. The carbonyl-containing nonhydrogen bonding groups (carbonyl, ester, peroxy acyl nitrate) have a lower a value of ~ 0.3.

The second order effects can be seen as modifications to the functional group contributions, and have likely steric and/or inductive causes. If a functional group is placed on a ring (as opposed to a chain), $\log_{10} p^0$ will be lower. On the other hand, if a functional group is placed not at or near the end of a chain (i.e. not at the 1 or 2 position) $\log_{10} p^0$ will be higher. As is well known, primary alcohols (i.e. where the hydroxyl is placed on a primary carbon) have lower vapour pressures than corresponding secondary alcohols, which in turn have lower vapour pressures than tertiary alcohols. The difference in $\log_{10} p^0$ is about the same between primary and secondary, and between secondary and tertiary alcohols. A double bond conjugated with a carbonyl functionality (aldehyde or ketone) lowers the vapour pressure. This is probably due to the increased dipole moment. p-values of the second order effects are all well below the 0.05 significance level.

For the hydrocarbons, there is an increase in MAD and predicted MAD compared to the regression for hydrocarbons only (see Sect. 4.2.1), but the performance is still satisfactory. For most molecule classes, MAD and pred. MAD are quite low, indicating the goodness-of-fit and the predictivity. The relatively lower performance of the model for peroxy acyl nitrates and peracids can be ascribed to the very limited number of molecules in the data set and possibly also to experimental uncertainty, as decomposition can be a problem for this type of molecules (Egerton et al., 1951; Kacmarek et al., 1978). The bad performance for peroxides, for which the number of data points seems acceptable, is more difficult to understand. Either their vapour pressures do not follow a simple group-contribution rule as for example for the ethers, or the data quality is particularly bad. The peroxide group, as the ether group, does not have a separate group contribution, as they are counted already in descriptor $k = 2$. Inserting a separate descriptor for peroxides did not improve significantly their performance.

We considered also some second order effects that are not retained in the final model. Apart from the p-value, also their influence on the pred. MAD was considered. In our previous method (Capouet and Müller, 2006), we distinguished between primary, secondary and tertiary nitrate groups. However, based on our current vapour pressure data set, we do not find this effect significant (p-value not below 0.05) and it is therefore not retained in the current method. On the other hand, the RI data of Fischer (1999); Kastler (1999) does suggest such an effect. More experimental vapour pressure data on nitrates will hopefully shed light on this issue.

Introducing a descriptor for branching next to peroxide groups (e.g. $-$C(C)OOC$-$), reduced the MAD from 0.39 to 0.25, but increased the pred. MAD from 0.40 to 0.51, and the p-value of this parameter was 0.08. Therefore, this descriptor was not retained. As opposed to carbonyl functionalities, no important impact was found for double bonds conjugated with acid or ester functionalities. Although branching next to hydrogen bonding groups (e.g. $-$C(C)C(=O)OH, $-$C(C)C(OH)$-$) seems to increase $\log_{10} p^0$(298K) by about 0.06 (p-value of 0.007), its impact on the MAD and pred. MAD of the hydrogen bonding compounds is marginal.

Table 2. Optimal parameter values for the model for hydrocarbons and monofunctional compounds (Sect. 4.2, Eqs. 13, 19 and 20).

Descriptor c_k	k	a_k	b_k	$g_k \pm \sigma_k$	
First order					
Zero-point ($c_1 = 1$)	1	2.62580	−1479.07	2.34 ± 0.02	
Size and topology					
# C + # in-chain O atoms[a]	2	0.06060	−2891.12	-0.501 ± 0.002	
t	3	−0.01887	1085.30	0.192 ± 0.006	
Functional groups					
# −ON(=O)=O	4	0.77005	−16 610.24	-2.46 ± 0.05	
# −C(=O)− (carbonyl[b])	5	0.31230	−7822.74	-1.21 ± 0.02	
# −C(=O)O−	6	0.29952	−4846.83	-0.64 ± 0.02	
# −C(=O)OON(=O)=O	7	0.35976	−15 609.26	-2.67 ± 0.19	
# −OH (hydroxyl)	8	0.89032	−16 027.87	-2.23 ± 0.02	
# −C(=O)OH	9	1.00026	−23 288.42	-3.53 ± 0.04	
# −OOH	10	0.85727	−19 313.13	-2.90 ± 0.09	
# −C(=O)OOH	11	0.82065	−18 337.43	-2.74 ± 0.13	
Second order					p-value
# X in ring[c]	12	−0.04758	−558.56	-0.16 ± 0.03	2×10^{-7}
# X on chain, and not at 1 or 2 position[c]	13	0.03569	182.52	0.07 ± 0.02	4×10^{-5}
# −C=C−C=O (carbonyl[b])	14	−0.12026	−736.37	-0.26 ± 0.04	9×10^{-7}
0, 1, 2 for prim., sec., tert. OH resp.	15	−0.02363	2132.47	0.39 ± 0.02	1×10^{-7}
1 if alkenoic alcohol, 0 otherwise	16	−0.32572	2823.11	0.22 ± 0.07	0.005

[a] Here and at following occurrences, "#" stands for "number of".
[b] Note that "carbonyl" designates ketone or aldehyde here, not e.g. ester, carboxylic acid.
[c] X = –**O**– (ether, ester), –**OO**–, –**C**ON(=O)=O, –**C**(=O)– (carbonyl, ester), –**C**(OH)–, –**C**(OOH). The location of the bold atom is considered.

4.3 Full method

4.3.1 Non-additivity in the A (or ΔS_b) term

An additive model as described above works well for non-substituted hydrocarbons and monofunctional compounds, but it breaks down in general for molecules with multiple functional groups, especially hydrogen bonding ones. For example, the vapour pressure of diols and diacids is lower than it would be expected from the purely additive model described in Sect. 4.2, with parameters from Table 2 (Fig. 3). To a smaller degree, this can also be the case for non-hydrogen bonding polar compounds, like diesters (Fig. 4).

To describe this nonadditive behaviour in $\log_{10} p^0$, we assume that while B can still be described as a sum over groups, A can be split up in three parts, and for two of them the group contributions a_k do not add linearly.

$$A = A_{\text{lin}} + A_{\text{CL}} + A_{\text{HB}} \quad (27)$$

$$A_{\text{lin}} = \sum_{k,\text{lin}} c_k a_k \quad (28)$$

$$A_{\text{CL}} = \sum_{k,\text{CL}} \frac{c_k a_k}{N_{\text{CL}}^r} \quad (29)$$

$$A_{\text{HB}} = \sum_{k,\text{H}} \frac{c_k a_k}{N_{\text{HB}}^r} \quad (30)$$

The first part (lin) contains groups that are additive: the groups needed to describe hydrocarbons ($k = 1-3$) and the nitrate group. CL (carbonyl-like) denotes groups with a C=O group that are not hydrogen bonding: carbonyls, esters and peroxy acyl nitrates. HB (hydrogen-bonding) includes the hydrogen bonding functionalities (hydroxyl, acid, hydroperoxide, and peracid functionalities). The optimal value of the exponent r must be between 0 (additivity of group contributions for A_{CL} and A_{HB}) and 1 (A_{CL} and A_{HB} are averages rather than sums of group contributions), N_{HB} is the number of hydrogen bonding functionalities and N_{CL} the total number of carbonyl, ester and peroxy acyl nitrate groups. Optimizing "by hand" resulted in an optimal value of $r = 0.5$. The non-additivity in A – or ΔS_b, see Eq. (18) – can be understood as follows: the higher molecular order in the liquid when introducing a second group (e.g. going from a monoalcohol to a diol) is smaller than for the first functional group (e.g. going from an alkane to an alcohol). So while Eqs. (29) and (30) are empirical, they can be thermodynamically rationalised. The value of r is assumed to be the same for A_{CL} and A_{HB}, but because of the smaller value of a_k (~ 0.3) for the CL group, the nonadditive behaviour is weaker than for A_{HB} ($a_k \sim 1.0$). Note that also in the vapour pressure formulation of Myrdal and Yalkowsky (1997) the entropy of boiling increases less than linearly with the number of hydrogen

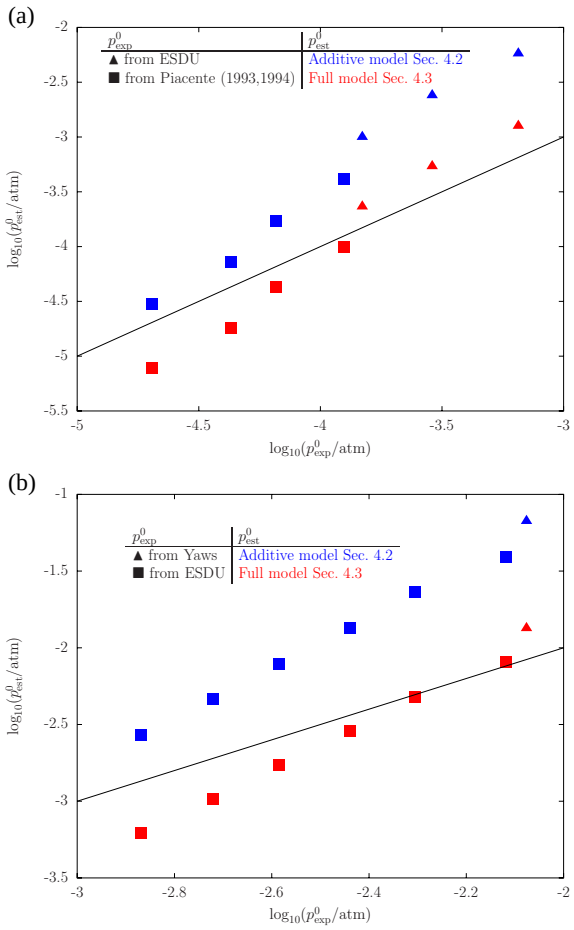

Fig. 3. (a): modeled (est) vs. experimental (exp) $\log_{10} p^0$ for linear α,ω-diols at 350 K, from butanediol to decanediol. Blue symbols: modeled results are obtained with the additive model of Sect. 4.2. This model, optimised for hydrocarbons and monofunctional compounds, overestimates the vapour pressure. Red symbols: modeled results are obtained with the full model of Sect. 4.3. (b): idem but for α,ω-diacids at 460 K. Smaller chains (2–3 carbon atoms) are not shown as local group-group interaction effects would mask the general trend. The black line is the 1:1 line.

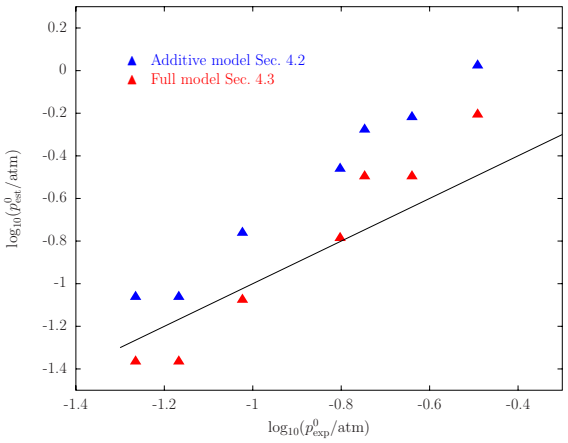

Fig. 4. Modeled (est) vs. experimental (exp) $\log_{10} p^0$ at 400 K for linear diesters, containing 4 to 8 carbon atoms. Blue symbols: modeled results are obtained with the additive model of Sect. 4.2. This model, optimised for hydrocarbons and monofunctional compounds, overestimates the vapour pressure. Red symbols: modeled results are obtained with the full model of Sect. 4.3. The black line is the 1:1 line.

bonding groups. By considering the separate terms A_{CL} and A_{HB}, one assumes additivity for CL and HB types of groups towards each other. This is supported by the data on hydroxy ketones.

As can be seen in Figs. 3 and 4, this approach prevents the systematic overestimation of vapour pressure for bifunctional compounds.

4.3.2 Modification for functionalised diacids

The recent data on functionalised diacids (e.g., Booth et al., 2010) points however to a much higher vapour pressure than predicted by the above formulation, and also higher than obtained by a simple group contribution method, with parameters obtained from less functionalised molecules. For example, citric acid, (6 carbon atoms, 3 acid functionalities, 1 hydroxyl functionality), has at 298 K a liquid vapour pressure that is higher by about an order of magnitude than that of adipic acid, (6 carbon atoms, 2 acid functionalities) (Booth et al., 2010), while values that are lower by roughly 5–6 orders of magnitude could be expected based on the simple group contribution method in Sect. 4.2. Likewise, according to the data of Booth et al. (2010), 2,3-dihydroxy succinic acid has a liquid vapour pressure that is higher by about two orders of magnitude than that of succinic acid, while the simple group contribution method would predict a value that is lower by roughly 3–4 orders of magnitude. This cannot be explained by uncertainties between references (see Sect. 2.5.3), as we took for these examples all vapour pressures from the same reference. Furthermore, we note that according to the high-temperature (above the melting point) liquid vapour pressure data of Yaws (1994), citric and tartaric acid have a lower vapour pressure than adipic acid and succinic acid, respectively, more in line with chemical intuition. A possible explanation could be that there are problems with the measurements of sublimation pressures of functionalised diacids, as Soonsin et al. (2010) had already concluded on the sublimation pressures of diacids. Unfortunately, as long as there are no other room-temperature sublimation pressure measurements available, this is difficult to verify. Including the non-additivity behaviour from the previous section would only increase the disagreement, since it tends to lower the modeled vapour pressure of a polyfunctional compound. Therefore, for this type of compounds (at least three CL

and/or HB functionalities, of which at least two acids) an effective group number N_{eff} is introduced:

$$c'_k = \frac{c_k}{N_{\text{CL}} + N_{\text{HB}}} N_{\text{eff}}, \text{ for } k \text{ of CL or HB type} \quad (31)$$

Optimizing "by hand" resulted in an optimal value of $N_{\text{eff}} = 2.6$. In contrast with the non-additivity behaviour discussed in the previous section, we cannot give a straightforward explanation of this behaviour, except that seemingly in such heavily functionalised molecules not all functional groups can bond efficiently in the liquid phase at the same time.

For polyols for example, it is seemingly not necessary to introduce this modification. An explanation could be that for the polyols in our data base (mostly with a linear carbon skeleton) efficient intermolecular interaction is possible despite the fact that they are heavily functionalised. Another explanation could lie in the fact that most data for polyols was obtained at high temperature (above, or closely below, the melting point), and that the need for Eq. (31) for functionalised diacids would be less for high temperature data.

4.3.3 Including intramolecular group-interactions, evaluation of all molecules

The set of descriptors and associated parameters of the full model is given in Table 4. These parameters are fitted to the data of all compounds (hydrocarbons, monofunctional and polyfunctional molecules).

The second order effect "X on chain and not at 1 or 2 position", which was still present for the method described in Sect. 4.2, was not retained here, as its g_k became very small (~ 0.01) and its p-value became very large (~ 0.5).

Except the "alkenoic alcohol flag", which is to be counted at most once per molecule, the other second order effects are counted for each functional group they apply to. Therefore, if applicable, groups 16 and 17 are to be counted twice for dicarbonyls (once for each ketone or aldehyde functionality) and once for carbonyl esters (once for the carbonyl functionality, but not for the ester functionality). A molecule with two carbonyl groups will have a higher vapour pressure if these groups are vicinal than if they are at more distant locations in the molecule, because the dipole moments of both carbonyl groups tend to cancel each other. Except for 2,3-butanedione, there are no room-temperature vapour pressures available of molecules with this structure, and this could be a reason why previous estimation methods did not take this effect into account. It can be illustrated with boiling points and with RI (Table 5). Both properties indicate that molecules with vicinal carbonyl functionalities have a significantly higher volatility than isomers with non-vicinal functionalities.

For diesters, we cannot discern a similar effect from the vapour pressure data. This is probably due to the lower dipole moment of an ester functionality. If a carbonyl group (ketone, aldehyde) is in β-position vs. another carbonyl group, this also leads to a higher vapour pressure.

Table 3. Evaluation of the model for hydrocarbons and monofunctional compounds (Sect. 4.2, Eqs. 13, 19 and 20).

	# Molecules	MAD[b]	Pred. MAD[c]
Hydrocarbons	130	0.073	0.075
Nitrates	23	0.087	0.095
Carbonyls[a]	128	0.064	0.069
Ethers	52	0.076	0.078
Esters	53	0.043	0.044
Peroxides	11	0.389	0.404
Peroxy acyl nitrates	5	0.106	0.226
Alcohols	120	0.076	0.080
Acids	49	0.077	0.080
Hydroperoxides	4	0.041	0.099
Peracids	4	0.215	0.304
All	579	0.071	0.075

[a] "Carbonyls" designate aldehyde or ketone here.
[b] Mean absolute deviation (Eq. 10) and [c] predicted mean absolute deviation (Eq. 12).

This is ascribed to the keto-enol tautomerism, an effect well-known in organic chemistry (e.g., Burdett and Rogers, 1964), where the keto-form is transformed into the less polar, more volatile, enolic form.

Intramolecular hydrogen-bonding for diols with vicinal hydroxyl functionalities leads to an increased vapour pressure and lower vaporisation enthalpy, as noted by Verevkin (2004). We noticed also for hydroxy carbonyls and hydroxy ethers an increase in vapour pressure if the two functionalities are vicinal. For hydroxy nitrates, direct vapour pressure data (see Roberts (1990) for a compilation) is very sparse (often only a single vapour pressure point) and of questionable accuracy, as vapour pressure was not the target property of the respective studies. However, from the RI on hydroxy nitrates of Kastler (1999) (see Table 6) a decrease in RI (increased p^0) is observed if both functionalities are vicinal. Therefore, we introduced one single descriptor for a functional group next to an hydroxyl functionality, leading to an increase in vapour pressure. Data on oxo acids is sparse, but it could nevertheless be concluded that p^0 increases if carbonyl and acid functionalities are vicinal. For hydroxy acids, no firm conclusions could be drawn in this respect.

We note that the RI data from Kastler (1999); Fischer (1999) on dinitrates suggests that the vicinality of nitrate functionalities also lead to an increase in p^0. The direct vapour pressure data does not allow to draw this conclusion, so this effect was not retained in the final model.

For hydrocarbons and monofunctional compounds together, the MAD and pred. MAD are 0.085 and 0.087 respectively. This is only slightly higher than as obtained with the method of Sect. 4.2 (see Table 3). For bifunctional compounds, the model works reasonably well, but with a lower performance for dinitrates, diacids, keto acids and hydroxy nitrates (Table 7). This can at least in part be ascribed to

Table 4. Descriptors and parameters for the full method (Sect. 4.3, Eqs. 13, 20 and 27–31).

Descriptor c_k	k	Type[d]	a_k	b_k	$g_k \pm \sigma_k$	
First order						
Zero-point ($c_1 = 1$)	1	lin	2.6255	−1986.56	2.24 ± 0.03	
Size and topology						
# C + # in-chain O atoms	2	lin	0.06298	−2821.46	−0.485 ± 0.004	
t	3	lin	−0.00293	1040.69	0.20 ± 0.01	
Functional groups						
# −ON(=O)=O	4	lin	0.71114	−15 841.13	−2.37 ± 0.03	
# −C(=O)− (carbonyl[e])	5	CL	0.19747	−7163.72	−1.20 ± 0.02	
# −C(=O)O−	6	CL	0.32257	−5208.53	−0.69 ± 0.03	
# −C(=O)OON(=O)=O	7	CL	0.29030	−15 011.33	−2.63 ± 0.33	
# −OH(hydroxyl)	8	HB	0.95537	−16 699.73	−2.29 ± 0.02	
# −C(=O)OH	9	HB	0.98567	−23 671.00	−3.62 ± 0.02	
# −OOH	10	HB	0.78348	−18 583.48	−2.83 ± 0.16	
# −C(=O)OOH	11	HB	0.81498	−18 071.50	−2.70 ± 0.24	
Second order						p-value
# X on ring[a]	12	lin, CL, HB[b]	0.18704	−2509.37	−0.30 ± 0.03	1×10^{-7}
# −C=C−C=O	13	CL	−0.18596	14.21	−0.18 ± 0.07	0.01
0, 1, 2 for prim., sec., tert. OH respectively	14	HB	−0.28012	4201.34	0.54 ± 0.03	1×10^{-7}
1 if alkenoic alcohol, 0 otherwise	15	HB	−0.34191	2961.95	0.23 ± 0.11	0.05
Intramolecular group interactions[c] Per carbonyl[e] group:						
# C=O (CL type) at α-position present	16	CL	0.26830	1602.62	0.58 ± 0.13	8×10^{-4}
# C=O (CL type) at β-position present[f]	17	CL	0.11716	939.71	0.30 ± 0.12	0.03
# Functional groups (not CL type nor acid) at α-position present	18	CL	−0.30373	3769.87	0.43 ± 0.13	0.01
Per hydroxyl group: # functional groups at α-position present	19	HB	−0.04143	800.12	0.11 ± 0.02	2×10^{-5}
Per acid group: # C=O (CL type) at α-position present	20	HB	0.46023	1817.69	0.81 ± 0.08	4×10^{-7}

[a] For X the same definitions as in Table 2 are applicable.
[b] The type depends on the type of functional group contribution to which this second order effect is applicable.
[c] Functional group at α-position with respect to another functional group means that both functional groups are vicinal: they are bonded to two adjacent carbon atoms. Examples: $-C(=O)C(=O)-$, $-CH(OH)CH_2OCH_2-$. Functional group at β-position with respect to another functional group means that they are bonded to two carbon atoms that are separated by one carbon atom. Example: $-C(=O)CH_2C(=O)-$.
[d] lin: additive groups. CL: carbonyl-like groups. HB: hydrogen-bonding groups.
[e] Note that 'carbonyl' designates ketone or aldehyde here, while with 'carbonyl-like', also esters and peroxy acyl nitrates are meant.
[f] This effect is not counted if, on the carbon in between, another functional group is present. It is hence counted for $-C(=O)CH_2C(=O)-$ but not for $-C(=O)CHOHC(=O)-$.

the experimental data, which is sparse and/or conflicting. Given that other bifunctional compounds are relatively well described, new experimental data can probably reduce these errors significantly, by updating the parameters but without having to modify the model framework. For compound classes with more than two functional groups, the pred. MAD can be very large, up to 0.69. Note that for molecule classes with only few compounds, the reported uncertainties are uncertain themselves. We checked for each vicinal group interaction descriptor, that its removal led to a significantly higher MAD and pred. MAD for some molecule classes.

Application examples of EVAPORATION for hydrocarbons, mono- and polyfunctional compounds are given in the Supplement.

5 Comparison with other methods

The considered methods are SPARC (SPARC performs automated reasoning in chemistry) version 4.2 (http://archemcalc.com/sparc) (Hilal et al., 2003), SIMPOL (SIMplified p_L^0 prediction method, Pankow and Asher, 2008), and the methods of Capouet and Müller (2006) (CM), of Myrdal and Yalkowsky (1997) (MY), and of Nannoolal et al. (2008) (Nan). The last two methods are combined with the boiling point methods of Joback and Reid (1987) (JR) or Nannoolal et al. (2004) (Nan): MY-Nan, MY-JR and Nan-Nan. These methods were already intercompared by Compernolle et al. (2010). Note that some of the original methods had to be extended (Compernolle et al., 2010) to treat certain functional groups (i.e. hydroperoxides, peracids). A short description of the methods is given in Table 8. We did not implement

Table 5. Boiling points and isothermal Kovats retention indices (RI) from GC on nonpolar columns for diones, both retrieved from NIST. This list is not meant to be complete, but only serves to illustrate the intramolecular effect of vicinal functional groups.

Molecule	T_b/K	RI
2,3-Pentanedione	384.2	653–675
2,4-Pentanedione	411	763–791
2,3-Hexanedione	401.2	755–764
3,4-Hexanedione	403.2	773
2,4-Hexanedione	431.2	800–900
2,5-Hexanedione	467.15	906
2,3-Octanedione	n.a.	968
2,4-Octanedione	n.a.	1079–1091

Table 6. Kovats retention indices (RI) from temperature-programmed GC from Kastler (1999) for hydroxy nitrates. This list is not meant to be complete, but only serves to illustrate the intramolecular effect of vicinal nitrate and hydroxy groups.

Molecule	RI
1-Hydroxypropyl-2-nitrate	818
1-Hydroxypropyl-3-nitrate	884
1-Hydroxybutyl-2-nitrate	915
4-Hydroxybutyl-2-nitrate	932
1-Hydroxyhexyl-2-nitrate	1111
6-Hydroxyhexyl-1-nitrate	1217

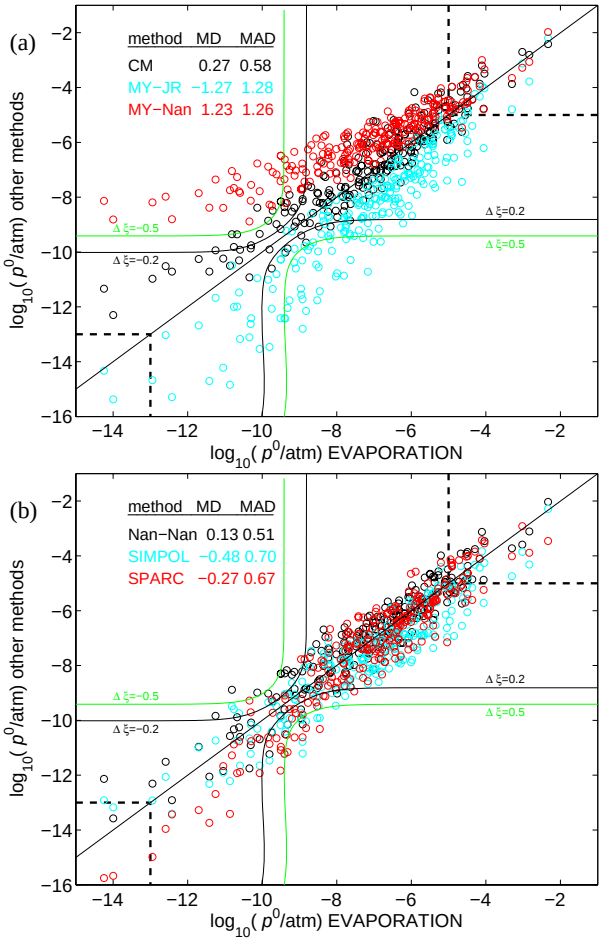

Fig. 5. $\log_{10}\frac{p^0}{\text{atm}}$ of various methods vs. that of EVAPORATION (the full method described in Sect. 4.3), for the explicit molecules in the chemical mechanism BOREAM for α-pinene degradation. **(a)**: the methods CM (black), MY-JR (blue) and MY-Nan (red). **(b)**: the methods Nan-Nan (black), SIMPOL (blue) and SPARC (red). Also given are the mean deviation (MD) and mean absolute deviation (MAD) of those methods vs. EVAPORATION. The black line is the 1:1 line. Data points above the dashed upper right corner indicate species with $p^0 > 10^{-5}$ atm that will not partition appreciable to the aerosol phase even at high aerosol loadings, while data points below the dashed lower left corner indicate species with $p^0 < 10^{-13}$ atm that will be almost exclusively in the aerosol phase, even at low aerosol loadings. Also shown are lines representing the vapour pressure needed to cause a change $\Delta\xi$ of ± 0.2 and ± 0.5 in the condensed fraction ξ, where ξ is calculated by Eq. (32) with $\overline{M}_{\text{aer}} = 200\,\text{g mol}^{-1}$, $\gamma_i = 1$, $T = 298\,\text{K}$, $C_{\text{aer}} = 3.16\,\mu\text{g m}^{-3}$ and p^0 provided by EVAPORATION.

the code of SPARC, as we do not have access to its current version, but we have calculated the vapour pressure of all condensable explicit species occurring in BOREAM on-line with SPARC, version 4.2.

5.1 Comparison of predicted vapour pressures for SOA compounds without experimental data

Figure 5 compares the $\log_{10}\frac{p^0}{\text{atm}}$ of various estimation methods vs. that of EVAPORATION (the full model of Sect. 4.3), which is taken as the base case. Intercomparing different estimation methods cannot determine which method is the best for estimating vapour pressures of SOA components, but it might help modellers to figure out the possible impact of using EVAPORATION on simulated aerosol yield. As in Compernolle et al. (2010), the test molecules are the explicit molecules present in the chemical mechanism BOREAM for α-pinene degradation by OH, O_3 and NO_3 (Capouet et al., 2008). Given are the MD and mean absolute deviation MAD of these methods vs. the base case. On average, EVAPORATION calculates somewhat lower vapour pressures than the CM method used in our previous modelling studies, indicating that simulated SOA yields will be higher upon implementation of this new method in the BOREAM model. However, currently no functionalised diacids are present in the explicit part of the BOREAM mechanism. Therefore, the empirical modification of the method, described in Sect. 4.3.2, does not play any role. Applying EVAPORATION to the α-pinene tracer 3-methyl-1,2,3-butane-

Table 7. Evaluation of the model for all compounds (Sect. 4.3, Eqs. 13, 20 and 27–31).

Molecule class	# Molec.	MD	Pred. MD	MAD	Pred. MAD
Hydrocarbons	130	0.033	0.032	0.094	0.096
Monofunctional					
Nitrates	23	2×10^{-4}	2×10^{-4}	0.072	0.073
Carbonyls[a]	128	−0.046	−0.047	0.084	0.087
Ethers	52	−0.007	−0.007	0.080	0.082
Esters	53	−0.020	−0.020	0.052	0.053
Peroxides	11	−0.273	−0.282	0.319	0.328
Peroxy acyl nitr.	5	0.022	−0.031	0.096	0.215
Alcohols	120	0.019	0.020	0.087	0.090
Acids	49	−0.054	−0.055	0.086	0.087
Hydroperoxides	4	−0.013	−0.046	0.044	0.065
Peracids	4	−0.039	−0.057	0.208	0.294
Bifunctional					
Dinitrates	10	0.088	0.092	0.281	0.294
Dicarbonyls[a]	18	−0.035	−0.064	0.072	0.106
Diols	32	0.027	0.029	0.124	0.131
Diacids	31	0.043	0.046	0.200	0.207
Diethers	16	−0.009	−0.010	0.110	0.112
Diesters	13	0.048	0.050	0.105	0.109
Carbonyl esters[a]	12	0.056	0.059	0.075	0.077
Carbonyl acids[a]	18	0.135	0.149	0.225	0.243
Hydroxy ethers	11	−0.038	−0.039	0.081	0.082
Hydroxy nitrates	4	0.333	0.339	0.333	0.339
Hydroxy peroxides	1	−0.397	−0.401	0.397	0.401
Hydroxy carbonyls[a]	17	0.040	0.045	0.116	0.129
Hydroxy acids	10	−0.032	−0.033	0.154	0.157
> 2 Functionalities					
Polyols	12	−0.008	−0.019	0.307	0.337
Polynitrates	5	−0.289	−0.661	0.328	0.691
At least 2 acids	12	−0.028	−0.019	0.404	0.434
Other	5	0.376	0.628	0.385	0.639
All	788	0.003	0.001	0.101	0.106

[a] "Carbonyl" designates aldehyde or ketone here.

tricarboxylic acid (MBTCA), found in substantial amounts in ambient aerosols (Szmigielski et al., 2007), a relatively high vapour pressure ($\sim 10^{-11}$ atm) is predicted, as compared to the other methods, except MY-Nan and SPARC (Table 9).

5.2 Sensitivity of partitioning to vapour pressure estimation

The major use of the knowledge of the vapour pressure of a compound is to estimate its tendency to partition into the particulate phase. The condensed fraction ξ of a compound i can be expressed as (e.g., Donahue et al., 2009; Valorso et al., 2011)

$$\xi_i^{\text{aer}} = 1/\left(1+\frac{\overline{M}_{\text{aer}}\gamma_i p_i^0}{C_{\text{aer}}RT}\right) \quad (32)$$

with $\overline{M}_{\text{aer}}$ the mean organic aerosol mass, γ_i the activity coefficient of the compound and C_{aer} the total organic aerosol mass concentration. C_{aer} varies typically between 0.1 (low aerosol loading) and 100 (high aerosol loading) $\mu g\,m^{-3}$. In analogy with Valorso et al. (2011), in Fig. 5 the regions below 10^{-13} atm and above 10^{-5} atm are indicated with dashed corners. Compounds with vapour pressures below 10^{-13} atm (above 10^{-5} atm) will be almost exclusively in the aerosol phase (gas phase) even for low (high) aerosol loadings.

Errors in p^0 will affect the condensed fraction ξ. Take a scenario where $\overline{M}_{\text{aer}} = 200\,\text{gmol}^{-1}$, $\gamma_i = 1$ (ideality assumption), $T = 298\,\text{K}$ and $C_{\text{aer}} = 3.16\,\mu g\,m^{-3}$ (the geometric mean of 0.1 and 100 $\mu g\,m^{-3}$). For this scenario, in Fig. 5 the change in $\log_{10}(p^0)$ is depicted that is needed to change ξ by $\Delta\xi = \pm 0.2$ and $\Delta\xi = \pm 0.5$. Errors in ξ

Table 8. Methods used in the intercomparison with EVAPORATION.

Method	Reference	Property	Short description
	p^0 estimation from molecular structure only		
CM	Capouet and Müller (2006)	p^0	Group contribution method[a]
SIMPOL	Pankow and Asher (2008)	p^0	Group contribution method
SPARC	Hilal et al. (2003)	p^0	Includes molecular descriptors (e.g. polarizability, dipole), which are themselves calculated from atomic fragments
	p^0 estimation from molecular structure and T_b		
MY	Myrdal and Yalkowsky (1997)	p^0	Includes two descriptors: for molecular flexibility and #hydrogen bonding groups
Nan	Nannoolal et al. (2008)	p^0	Detailed group contribution method
	T_b estimation from molecular structure		
Nan	Nannoolal et al. (2004)		Detailed group contribution method
JR	Joback and Reid (1987)		Group contribution method

[a] The parent hydrocarbon part is estimated by combining work from Marrero and Gani (2001) (MG) and Ambrose and Walton (1989) (AW) (for details see Compernolle et al., 2010).

Table 9. $\log_{10}(p^0/\text{atm})$ of 3-methyl-1,2,3-butane-tricarboxylic acid (MBTCA), as calculated by various methods including EVAPORATION. No experimental value is available.

Method	$\log_{10}\left(p^0/\text{atm}\right)$
CM	−11.14
MY-JR	−13.08
MY-Nan	−9.26
Nan-Nan	−12.05
SIMPOL	−12.09
SPARC	−10.61
EVAPORATION	−10.86

of 0.2 and 0.5 can be considered significant and grave respectively. At $p^0 = p^* \equiv C_{\text{aer}} RT/\overline{M}_{\text{aer}}$ ($3.87 \cdot 10^{-10}$ atm for this scenario), ξ is most sensitive to changes in p^0. This sensitivity decreases above and below p^*. Note that considering another scenario (with another p^*) would shift the curve along the diagional but would preserve the shape. If the variation between MY-Nan and MY-JR -giving the highest and lowest vapour pressures, respectively, of all methods considered here- would reflect the uncertainty in vapour pressure estimation, the errors in estimating the condensed fraction ξ would be grave (i.e. $|\Delta\xi| > 0.5$) even orders of magnitude above and below p^*. As compared to EVAPORATION, MY-JR gives grave overestimations of ξ up to about $1.6 \cdot 10^{-7}$ atm (a factor 400 above p^*) and significant overestimations ($|\Delta\xi| > 0.2$) up to about -6 atm (a factor 2600 above p^*). For the low vapour pressures, the situation is even worse: as compared to EVAPORATION, MY-Nan gives grave underestimations of ξ for the entire given range below p^*.

However, this view is probably too pessimistic. The study of Barley and McFiggans (2010), applied to relatively volatile compounds (as compared to typical OA components), showed that MY-JR and MY-Nan under- and overestimates experimental vapour pressures, respectively. This discrepancy is likely to increase for compounds with lower p^0. If MY-JR and MY-Nan are omitted, grave errors occur up to a factor 25 above p^* (underestimation, compared to EVAPORATION, by SIMPOL) and a factor 25 below p^* (overestimation, compared to EVAPORATION, by CM). Significant errors occur up to a factor 320 above p^* (underestimation, compared to EVAPORATION, by SIMPOL) and a factor 70 below p^* (overestimation, compared to EVAPORATION, by CM). Hence in a region of 1–2 orders of magnitude around p^*, significant errors in ξ estimation can be expected. If an error of $|\Delta\xi| < 0.2$ is desired, even at p^*, one calculates from Eq. (32) that the error on $\log_{10} p^0$ should be below 0.37. While the pred. MAD of EVAPORATION is below this treshold for hydrocarbons, most monofunctional and bifunctional classes (see Table 7), this is not the case for most molecules classes with more than two functional groups. To get the error on $\log_{10} p^0$ below 0.37 also for polyfunctional compounds is a major challenge.

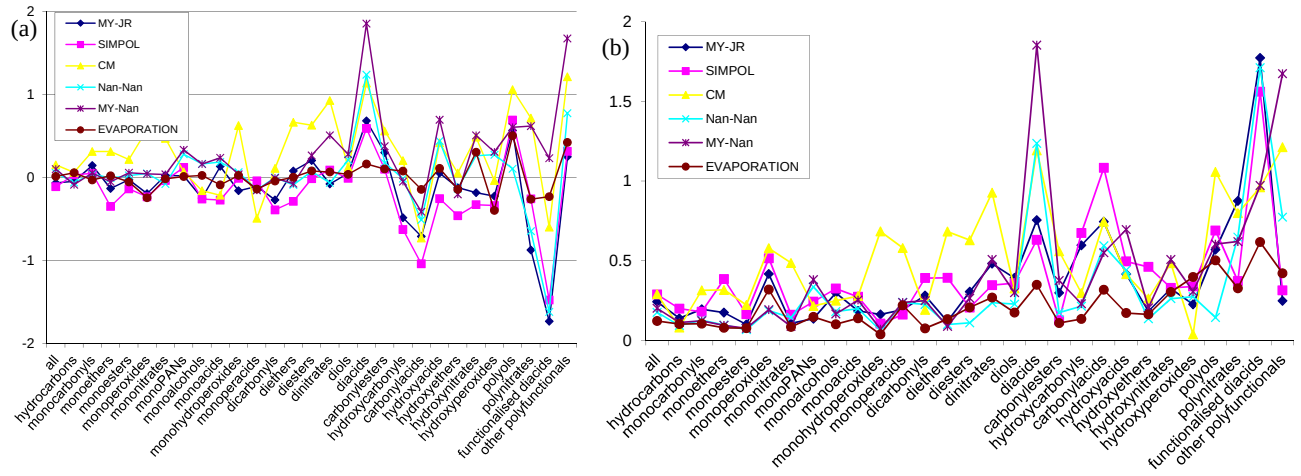

Fig. 6. MD (a) and MAD (b) of various vapour pressure estimation methods, including EVAPORATION (the full method described in Sect. 4.3) for all compounds (first point) and for different molecule classes, with experimental data points selected between 270 and 390 K from our data base. See text for details.

5.3 Comparison with experimental data points

We have also compared the various methods to our experimental data set of vapour pressures. While EVAPORATION was fitted also for high temperatures (up to the critical temperature if available) this is not the case for most other methods, and it would be unreasonable to test them for these high temperatures. On the other hand, the restriction to atmospherically relevant temperatures (say up to 40 °C) would leave out several molecule classes (e.g. most polyols). Therefore, we took a temperature range of 270 to 390 K. Another requirement was that the temperature had to be below the critical temperature as estimated by the method of Marrero and Gani (2001) (MG) of the parent hydrocarbon, as otherwise the CM method (Capouet and Müller, 2006) would fail. SPARC was not considered in this intercomparison, as the number of vapour pressure points was too high to calculate by this on-line method. Figure 6 summarizes the MD and MAD for all methods for different molecule classes. One can conclude that

– Even for monofunctional compounds, the CM method shows larger deviations. The main reason is that the CM method was optimised only within 298–320 K, a much narrower range than the considered temperature interval of 270–390 K.

– The MY-Nan method follows closely the Nan-Nan method for hydrocarbons and monofunctional compounds but diverges for more functionalised compounds, for which MY-Nan generally predicts higher vapour pressures than Nan-Nan. This is logical as both methods use the same boiling point method of Nannoolal et al. (2004), and the difference is evident only when the temperature is well below the boiling point. For some molecule classes, the overestimation of MY-Nan is extreme, reaching almost two orders of magnitude.

– SIMPOL and MY-JR show some of the largest underestimations.

– The largest deviations are seen for diacids, carbonyl acids, functionalised diacids and the rest group "other polyfunctionals" (see Supplement for their identity). Most methods overestimate diacid vapour pressure and underestimate vapour pressure of oxo acids and functionalised diacids. We note here that for diacids, Pankow and Asher (2008) selected some data (e.g., Chattopadhyay and Ziemann, 2005) that we chose not to include (see Sect. 2.5.3 for the motivation).

– EVAPORATION shows generally the lowest deviations. This is of course not a surprise as EVAPORATION was fitted to the data. We note however that for most molecule classes, the predicted MAD of EVAPORATION is only slightly above the MAD from the fitting (see Sect. 4.3.3).

6 Conclusions

A new vapour pressure estimation method has been developed, EVAPORATION, intended for polyfunctional molecules as they occur in SOA. Important features are the non-additivity in $\log_{10} p^0$ of functional groups, especially hydrogen-bonding ones, intramolecular group interactions, and the inclusion of recent data on functionalised diacids. To describe this last type of compounds, a modification had

to be introduced, effectively limiting the number of groups which are taken into account. We cannot provide a straightforward explanation for this behaviour. Although there is less data on functionalised diacids than on diacids, it is also in this case clear that important differences exist between different reference sources. E.g. sublimation pressure data for 2-oxoglutaric acid can differ by almost two orders of magnitude between different reference sources (Booth et al., 2010; Chattopadhyay and Ziemann, 2005; Frosch et al., 2010). If the experimental methodology of Soonsin et al. (2010), with the use of mixtures with water, were applied to obtain vapour pressures of functionalised diacids, the divergence would likely increase, as their sublimation pressure data for diacids is the lowest available. Counter-intuitively, the subcooled liquid data -calculated from sublimation pressures- of Booth et al. (2010) suggest that citric and tartaric acid have higher vapour pressures than adipic and succinic acid respectively although they have more polar groups. The high-temperature (above the melting point) liquid vapour pressure data of Yaws (1994) however suggest the reverse. Moreover, the data of Yaws (1994) on functionalised diacids, and of ESDU on diacids, after bold extrapolation, does obey the ΔH_v vs. $\log_{10}(p^0)$ correlation at 298 K already established for a wide variety of compounds (MacLeod et al., 2007; Epstein et al., 2010). Most ambient temperature data on diacids and functionalised diacids does not (except e.g. Cappa et al., 2007). One possible explanation is that there are problems with the measurements. More light can hopefully be shed on this issue by the measurement of high-temperature (above the melting point) liquid vapour pressure of other functionalised diacids. Also confirmation of the data of Yaws (1994) is desired, as this is a secondary source with no details on the experimental procedure. The high-temperature liquid vapour pressures should be relatively reliable: no solid to subcooled liquid (e.g., Booth et al., 2010) or mixture to pure liquid conversion (e.g., Riipinen et al., 2007) would be needed, and the vapour pressure should be more accurately measurable at this higher temperature. Also room temperature measurement of subcooled liquid vapour pressure of functionalised diacids by the methodology of Soonsin et al. (2010) can provide more insight, by comparing it with the existing solid vapour pressure data.

Vapour pressures of zero-, mono- and bifunctional compounds can be reasonably well predicted by EVAPORATION, while it performs worse for molecules with more functional groups but still better than other methods. This can at least in part be attributed to the fact that experimental error on the vapour pressures of these compounds is higher, evidenced by the disagreement between different reference sources in the case of diacids. On the other hand, it is to be expected that our -still relatively simple- model does not grasp completely the complex group-group interactions. However, to develop more detailed models, additional and more accurate data is a prerequisite. Within the present framework of EVAPORATION, better performance can be reached if more data is collected for these molecule classes with currently limited data availability (e.g. peracids, hydroxy nitrates).

Supplementary material related to this article is available online at: http://www.atmos-chem-phys.net/11/9431/2011/acp-11-9431-2011-supplement.pdf.

Acknowledgements. This work was supported by the projects IBOOT (SD/AT/03B, 2006–2010) and BIOSOA (SD/CS/05A, 2011-2014) funded by the Belgian Science Policy Office.

Edited by: B. Ervens

References

Aldrich: Catalog Handbook of Fine Chemicals, Aldrich Chemical Company, Inc., Milwaukee, WI, USA, 1990.

Ambrose, D. and Walton, J.: Vapour pressures up to their critical temperatures of normal alkanes and 1-alkanols, Pure Appl. Chem., 61, 1395–1403, 1989.

Asher, W. E. and Pankow, J. F.: Vapor pressure prediction for alkenoic and aromatic organic compounds by a UNIFAC-based group contribution method, Atmos. Environ., 40, 3588–3600, 2006.

Asher, W. E., Pankow, J. F., Erdakos, G. B., and Seinfeld, J. H.: Estimating the vapor pressures of multi-functional oxygen-containing organic compounds using group contribution methods, Atmos. Environ., 36, 1483–1498, 2002.

Aumont, B., Szopa, S., and Madronich, S.: Modelling the evolution of organic carbon during its gas-phase tropospheric oxidation: development of an explicit model based on a self generating approach, Atmos. Chem. Phys., 5, 2497–2517, doi:10.5194/acp-5-2497-2005, 2005.

Barley, M. H. and McFiggans, G.: The critical assessment of vapour pressure estimation methods for use in modelling the formation of atmospheric organic aerosol, Atmos. Chem. Phys., 10, 749–767, doi:10.5194/acp-10-749-2010, 2010.

Besalu, E.: Fast computation of cross-validated properties in full linear leave-many-out procedures, J. Math. Chem., 29, 191–204, 2001.

Bilde, M. and Pandis, S. N.: Evaporation rates and vapor pressures of individual aerosol species formed in the atmospheric oxidation of α- and β-pinene, Environ. Sci. Technol., 35, 3344–3349, 2001.

Bilde, M., Svenningsson, B., Monster, J., and Rosenorn, T.: Even-odd alternation of evaporation rates and vapor pressures of C3–C9 dicarboxylic acid aerosols, Environ. Sci. Technol., 37, 1371–1378, 2003.

Booth, A. M., Barley, M. H., Topping, D. O., McFiggans, G., Garforth, A., and Percival, C. J.: Solid state and sub-cooled liquid vapour pressures of substituted dicarboxylic acids using Knudsen Effusion Mass Spectrometry (KEMS) and Differential

Scanning Calorimetry, Atmos. Chem. Phys., 10, 4879–4892, doi:10.5194/acp-10-4879-2010, 2010.

Booth, A. M., Montague, W. J., Barley, M. H., Topping, D. O., McFiggans, G., Garforth, A., and Percival, C. J.: Solid state and subcooled liquid vapour pressures of cyclic aliphatic dicarboxylic acids, Atmos. Chem. Phys., 11, 655–665, doi:10.5194/acp-11-655-2011, 2011.

Bruckmann, P. W. and Willner, H.: Infrared spectroscopic study of peroxyacetyl nitrate (PAN) and its decomposition products, Environ. Sci. Technol., 17, 352–357, 1983.

Burdett, J. L. and Rogers, M. T.: Keto-enol tautomerism in β-dicarbonyls studied by nuclear magnetic resonance spectroscopy.[1] I. Proton chemical shifts and equilibrium constants of pure compounds, J. Am. Chem. Soc., 86, 2105–2109, 1964.

Capouet, M. and Müller, J.-F.: A group contribution method for estimating the vapour pressures of α-pinene oxidation products, Atmos. Chem. Phys., 6, 1455–1467, doi:10.5194/acp-6-1455-2006, 2006.

Capouet, M., Müller, J.-F., Ceulemans, K., Compernolle, S., Vereecken, L., and Peeters, J.: Modeling aerosol formation in alpha-pinene photooxidation experiments, J. Geophys. Res., 113, D02308, 2008.

Cappa, C. D. and Wilson, K. R.: Evolution of organic aerosol mass spectra upon heating: implications for OA phase and partitioning behavior, Atmos. Chem. Phys., 11, 1895–1911, doi:10.5194/acp-11-1895-2011, 2011.

Cappa, C. D., Lovejoy, E. R., and Ravishankara, A. R.: Determination of evaporation rates and vapor pressures of very low volatility compounds: A study of the C-4–C-10 and C-12 dicarboxylic acids, J. Phys. Chem. A, 111, 3099–3109, 2007.

Ceulemans, K., Compernolle, S., Peeters, J., and Müller, J.-F.: Evaluation of a detailed model of secondary organic aerosol formation from [alpha]-pinene against dark ozonolysis experiments, Atmos. Environ., 40, 5434–5442, 2010.

Chattopadhyay, S. and Ziemann, P. J.: Vapor pressures of substituted and unsubstituted monocarboxylic and dicarboxylic acids measured using an improved thermal desorption particle beam mass spectrometry method, Aerosol Sci. Technol., 39, 1085–1100, 2005.

Compernolle, S., Ceulemans, K., and Müller, J.-F.: Technical Note: Vapor pressure estimation methods applied to secondary organic aerosol constituents from α-pinene oxidation: an intercomparison study, Atmos. Chem. Phys., 10, 6271–6282, doi:10.5194/acp-10-6271-2010, 2010.

Compernolle, S., Ceulemans, K., and Müller, J.-F.: Technical Note: Estimating fusion properties for polyacids, Atmos. Chem. Phys. Discuss., 11, 7535–7553, doi:10.5194/acpd-11-7535-2011, 2011a.

Compernolle, S., Ceulemans, K., and Müller, J.-F.: Estimating fusion properties for functionalised acids, Atmos. Chem. Phys., 11, 8385–8394, doi:10.5194/acp-11-8385-2011, 2011b.

Donahue, N. M., Robinson, A. L., and Pandis, S. N.: Atmospheric organic particulate matter: From smoke to secondary organic aerosol, Atmos. Environ., 43, 94–106, 2009.

Dykyj, J., Svoboda, J., Wilhoit, R. C., Frenkel, M., and Hall, K. R.: Vapor Pressure and Antoine Constants for Hydrocarbons, and S, Se, Te, and Halogen Containing Organic Compounds Landolt-Bernstein: Numerical Data and Functional Relationships in Science and Technology, 20A, Springer, Berlin, Heidelberg, Germany, 1999.

Egerton, A. C., Emte, W., and Minkoff, G. J.: Some properties of organic peroxides, Discuss. Faraday Soc., 10, 278–282, 1951.

Epstein, S. A., Riipinen, I., and Donahue, N. M.: A semiempirical correlation between enthalpy of vaporization and saturation concentration for organic aerosol, Environ. Sci. Technol., 44, 743–748, 2010.

ESDU: Engineering Sciences Data Unit Ltd: Vapour pressures and critical points of liquids, ESDU International, 1995, 2001.

Everitt, B. S.: The Cambridge Dictionary of Statistics, Cambridge University Press, Cambridge, UK, 4th edn., 127–128, 2010.

Fischer, R. G.: Vorkommen und Umweltverhalten der C1–C10 organischen Nitrate in kontinentaler und mariner Luft, Ph.D. thesis, Abteilung Analytische Chemie und Umweltchemie, Universität Ulm, http://vts.uni-ulm.de/docs/1999/355/vts_355.pdf, 1999.

Fischer, R. C., Wittlinger, R., and Ballschmiter, K.: Retention-index based vapor pressure estimation for polychlorobiphenyl (PCB) by gas chromatography, Fresenius J. Anal. Chem., 342, 421–425, 1992.

Frosch, M., Zardini, A. A., Platt, S. M., Müller, L., Reinnig, M.-C., Hoffmann, T., and Bilde, M.: Thermodynamic properties and cloud droplet activation of a series of oxo-acids, Atmos. Chem. Phys., 10, 5873–5890, doi:10.5194/acp-10-5873-2010, 2010.

Hilal, S. H., Karickhoff, S. W., and Carreira, L. A.: Prediction of the vapor pressure boiling point, heat of vaporization and diffusion coefficient of organic compounds, Qsar Comb. Sci., 22, 565–574, 2003.

Jenkin, M. E., Saunders, S. M., and Pilling, M. J.: The tropospheric degradation of volatile organic compounds: a protocol for mechanism development, Atmos. Environ., 31, 81–104, 1997.

Joback, K. and Reid, R.: Estimation of pure-component properties from group-contributions, Chem. Eng. Commun., 57, 233–243, 1987.

Kacmarek, A. J., Solomon, I. J., and Lustig, M.: Preparation and properties of peroxyacetyl nitrate, J. Inorg. Nucl. Chem., 40, 574–576, 1978.

Kames, J. and Schurath, U.: Henry's law and hydrolysis-rate constants for peroxyacyl nitrates (PANs) using a homogeneous gas-phase source, J. Atmos. Chem., 21, 151–164, 1995.

Kastler, J.: Analytik, Massenspektrometrie und Vorkommen multifunktioneller Alkylnitrate in belasteter und unbelasteter Atmosphäre, Ph.D. thesis, Abteilung Analytische Chemie und Umweltchemie, Universität Ulm, 1999.

KDB: Korean thermophysical properties Data Bank, http://www.cheric.org/kdb/, last access: 2009–2010.

Korsten, H.: Internally consistent prediction of vapor pressure and related properties, Ind. Eng. Chem. Res., 39, 813–820, 2000.

Lide, D. R. (Ed.): CRC Handbook of Chemistry and Physics, 80th edn., CRC Press, Boca Raton, FL, USA, 2000.

Linstrom, P. and Mallard, W. (Eds.): NIST Chemistry WebBook, NIST Standard Reference Database Number 69, National Institute of Standards and Technology, Gaithersburg MD, 20899, http://webbook.nist.gov/chemistry/, last access: 2009–2010.

MacLeod, M., Scheringer, M., and Hungerbuhler, K.: Estimating enthalpy of vaporization from vapor pressure using Trouton's rule, Environ. Sci. Technol., 41, 2827–2832, 2007.

Marrero, J. and Gani, R.: Group-contribution based estimation of pure component properties, Fluid Phase Equilibr., 183–184, 183–

208, 2001.

Miyamoto, S., Nakamura, S., Iwai, Y., and Arai, Y.: Measurement of vapor-phase compressibility factors of monocarboxylic acids using a flow-type apparatus and their association constants, J. Chem. Eng. Data, 44, 48–51, 1999.

Moller, B.: Addendum to the method of Moller et al., in preparation, 2011.

Monster, J., Rosenorn, T., Svenningsson, B., and Bilde, M.: Evaporation of methyl- and dimethyl-substituted malonic, succinic, glutaric and adipic acid particles at ambient temperatures, J. Aerosol Sci., 35, 1453–1465, 2004.

Moller, B., Rarey, J., and Ramjugernath, D.: Estimation of the vapour pressure of non-electrolyte organic compounds via group contributions and group interactions, J. Mol. Liq., 143, 52–63, 2008.

Myrdal, P. B. and Yalkowsky, S. H.: Estimating pure component vapor pressures of complex organic molecules, Ind. Eng. Chem. Res., 36, 2494–2499, 1997.

Nannoolal, Y., Rarey, J., Ramjugernath, D., and Cordes, W.: Estimation of pure component properties: Part 1. Estimation of the normal boiling point of non-electrolyte organic compounds via group contributions and group interactions, Fluid Phase Equilib., 226, 45–63, 2004.

Nannoolal, Y., Rarey, J., and Ramjugernath, D.: Estimation of pure component properties: Part 3. Estimation of the vapor pressure of non-electrolyte organic compounds via group contributions and group interactions, Fluid Phase Equilibr., 269, 117–133, 2008.

Pankow, J. F.: An absorption model of the gas/aerosol partitioning involved in the formation of secondary organic aerosol, Atmos. Environ., 28, 189–193, 1994.

Pankow, J. F. and Asher, W. E.: SIMPOL.1: a simple group contribution method for predicting vapor pressures and enthalpies of vaporization of multifunctional organic compounds, Atmos. Chem. Phys., 8, 2773–2796, doi:10.5194/acp-8-2773-2008, 2008.

Poling, B. E., Prausnitz, J. M., and O'Connell, J. P.: The properties of gases and liquids, McGraw-Hill, New York, USA, 5th edn., A.47–A.60, 2001.

Pope, F. D., Tong, H.-J., Dennis-Smither, B. J., Griffiths, P. T., Clegg, S. L., Reid, J. P., and Cox, R. A.: Studies of single aerosol particles containing malonic acid, glutaric acid, and their mixtures with sodium chloride. II. Liquid-state vapor pressures of the acids, J. Phys. Chem. A, 114, 10156–10165, 2010.

Prausnitz, J. M., Lichtenthaler, R. N., and de Azevedo, E. G.: Molecular Thermodynamics of Fluid-Phase Equilibria, 3rd edn., Prentice-Hall PTR, Upper Saddle River, New Jersey 07458, 1999.

Raventos-Duran, T., Camredon, M., Valorso, R., Mouchel-Vallon, C., and Aumont, B.: Structure-activity relationships to estimate the effective Henry's law constants of organics of atmospheric interest, Atmos. Chem. Phys., 10, 7643–7654, doi:10.5194/acp-10-7643-2010, 2010.

Ribeiro da Silva, M. A. V., Monte, M. J. S., and Ribeiro, J. R.: Vapour pressures and the enthalpies and entropies of sublimation of five dicarboxylic acids, J. Chem. Thermodyn., 31, 1093–1107, 1999.

Ribeiro da Silva, M. A. V., Monte, M. J. S., and Ribeiro, J. R.: Standard enthalpies, entropies, and Gibbs functions of sublimation of four alkyl-substituted malonic acids, J. Chem. Eng. Data, 45, 756–759, 2000.

Ribeiro da Silva, M. A. V., Monte, M. J. S., and Ribeiro, J. R.: Thermodynamic study on the sublimation of succinic acid and of methyl- and dimethyl-substituted succinic and glutaric acids, J. Chem. Thermodyn., 33, 23–31, 2001.

Riipinen, I., Koponen, I., Frank, G., Hyvarinen, A.-P., Vanhanen, J., Lihavainen, H., Lehtinen, K., Bilde, M., and Kulmala, M.: Adipic and malonic acid aqueous solutions: surface tensions and saturation vapor pressures, J. Phys. Chem. A, 111, 12995–13002, 2007.

Roberts, J. M.: The atmospheric chemistry of organic nitrates, Atmos. Environ., 24, 243–287, 1990.

Roux, M. V., Temprado, M., and Chickos, J. S.: Vaporization, fusion and sublimation enthalpies of the dicarboxylic acids from C4 to C14 and C16, J. Chem. Thermodyn., 37, 941–953, 2005.

Salo, K., Jonsson, Ã. M., Andersson, P. U., and Hallquist, M.: Aerosol volatility and enthalpy of sublimation of carboxylic acids, J. Phys. Chem. A, 114, 4586–4594, 2010.

Sanchez, J. and Myers, T. N.: Peroxides and Peroxide Compounds, Organic Peroxides, vol. 18 of Kirk-Othmer Encyclopedia of Chemical Technology, 4th edn., Wiley, New York, USA, 2000.

Sander, R.: Compilation of Henry's Law Constants for Inorganic and Organic Species of Potential Importance in Environmental Chemistry, Public Report, http://www.rolf-sander.net/henry/henry.pdf, 1999.

Saunders, S. M., Jenkin, M. E., Derwent, R. G., and Pilling, M. J.: Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds, Atmos. Chem. Phys., 3, 161–180, doi:10.5194/acp-3-161-2003, 2003.

Soonsin, V., Zardini, A. A., Marcolli, C., Zuend, A., and Krieger, U. K.: The vapor pressures and activities of dicarboxylic acids reconsidered: the impact of the physical state of the aerosol, Atmos. Chem. Phys., 10, 11753–11767, doi:10.5194/acp-10-11753-2010, 2010.

Stein, S. E. and Brown, R. L.: Estimation of normal boiling points from group contributions, J. Chem. Inf. Comput. Sci., 34, 581–587, 1994.

Szmigielski, R., Surratt, J. D., Gómez-González, Y., Van der Veken, P., Kourtchev, I., Vermeylen, R., Blockhuys, F., Jaoui, M., Kleindienst, T. E., Lewandowski, M., Offenberg, J. H., Edney, E. O., Seinfeld, J. H., Maenhaut, W., and Claeys, M.: 3-methyl-1,2,3-butanetricarboxylic acid: An atmospheric tracer for terpene secondary organic aerosol, Geophys. Res. Lett., 34, L24811, doi:10.1029/2007GL031338, 2007.

Valorso, R., Aumont, B., Camredon, M., Raventos-Duran, T., Mouchel-Vallon, C., Ng, N. L., Seinfeld, J. H., Lee-Taylor, J., and Madronich, S.: Explicit modelling of SOA formation from α-pinene photooxidation: sensitivity to vapour pressure estimation, Atmos. Chem. Phys., 11, 6895–6910, doi:10.5194/acp-11-6895-2011, 2011.

Verevkin, S. P.: Determination of vapor pressures and enthalpies of vaporization of 1,2-alkanediols, Fluid Phase Equilibr., 224, 23–29, 2004.

Weast, R. and Grasselli, J.: CRC Handbook of Data on Organic Compounds, CRC Press, Inc., Boca Raton, FL, 2nd edn., 1989.

Yaws, C. L.: Handbook of Vapor Pressure, Gulf Publishing Company, Houston, USA, 1994.

Zardini, A. A., Krieger, U. K., and Marcolli, C.: White light Mie resonance spectroscopy used to measure very low vapor ressures of substances in aqueous solution aerosol particles, Opt. Express, 14, 6951–6962, 2006.