

The critical assessment of vapour pressure estimation methods for use in modelling the formation of atmospheric organic aerosol

M. H. Barley and G. McFiggans

Centre for Atmospheric Sciences, School of Earth, Environmental and Atmospheric Science, University of Manchester, Manchester, UK

Received: 6 August 2009 – Published in Atmos. Chem. Phys. Discuss.: 3 September 2009

Revised: 5 January 2010 – Accepted: 6 January 2010 – Published: 25 January 2010

Abstract. A selection of models for estimating vapour pressures have been tested against experimental data for a set of compounds selected for their particular relevance to the formation of atmospheric aerosol by gas-liquid partitioning. The experimental vapour pressure data (all <100 Pa) of 45 multifunctional compounds provide a stringent test of the estimation techniques, with a recent complex group contribution method providing the best overall results. The effect of errors in vapour pressures upon the formation of organic aerosol by gas-liquid partitioning in an atmospherically relevant example is also investigated. The mass of organic aerosol formed under typical atmospheric conditions was found to be very sensitive to the variation in vapour pressure values typically present when comparing estimation methods.

1 Introduction

The atmospheric aerosol is important for the understanding of climate change and for human health. The formation of condensed organic aerosol (OA) component mass from volatile organic compounds (VOCs) in the atmosphere is frequently described by a gas-liquid partitioning model (Barley et al., 2009, and references therein, e.g. Pankow, 1994), in which vapour pressure is a primary determinant. Whether the goal is to model OA composition using explicit methods (e.g. Aumont et al., 2005 and Bloss et al., 2005) or by non-arbitrary selection of model compounds from real candidate compound ensembles, accurate estimates of the vapour pressures for all the thousands of organic compounds at ambient temperatures will be required.

Reliable experimental vapour pressure data for a wide range of compounds are available from established databases (e.g. Dortmund Databank (DDB); http://www.ddbst.de/new/frame_DDB.htm). However the vast majority of these data have been collected by or on behalf of the chemical industry for chemical plant design improvement, with a particular emphasis on production and purification of products by distillation. Most of the data are therefore collected for structurally simple compounds (particularly hydrocarbons) with intermediate vapour pressures (10^3 – 10^5 Pa) and few data are collected at pressures <1 Pa. Most of the compounds found in the atmosphere will not have vapour pressure data available in the standard databases. Those atmospheric compounds most likely to condense into an aerosol will be relatively complex, high molecular weight (150–300) multifunctional compounds (Saxena and Hildemann, 1996) with ambient vapour pressures <0.1 Pa (Barsanti and Pankow, 2004); in many cases, orders of magnitude below 0.1 Pa. These compounds are very poorly represented in collections of experimental vapour pressure data and there are very few organic compounds for which experimental vapour pressure values have been measured below 0.01 Pa.

The aim of the current work is to evaluate a number of vapour pressure estimation techniques against those reliable primary data of most relevance to prediction of vapour pressures of likely OA components. The sensitivity of OA formation to variation in estimated vapour pressure values will also be investigated

1.1 Experimental measurement of vapour pressures

The accurate experimental measurement of low (1 – 10^3 Pa) and very low (<1 Pa) vapour pressures is a significant challenge. The use of modern pressure gauges means that it is theoretically possible to use the static method down to very low pressures but adsorption of volatiles (especially water) onto the surface of the apparatus and the presence of



Correspondence to: G. McFiggans
(g.mcfiggans@manchester.ac.uk)

impurities in the sample make this method difficult to use in practice at low pressures. Well established methods that do provide good results at low pressures are the transpiration method (Verevkin et al., 2000) and Knudsen effusion (Dekruif and Vanginkel, 1977; Hallquist et al., 1997). In the transpiration method a stream of inert gas (usually nitrogen) is used to slowly purge the headspace above a sample of the compound. The material picked up by the stream of nitrogen is then trapped (usually cryogenically) and determined after several hours purging. From the mass of compound transferred by the flowing gas stream the vapour pressure can be calculated. Good results are claimed for this method down to 0.005 Pa (Verevkin et al., 2000). In Knudsen effusion a small sample is placed in a closed container with a small hole in the side. The vapour in equilibrium with the sample (which is often a solid) effuses out of the small hole into a vacuum. The rate of mass loss allows the calculation of the vapour pressure. For solid samples the experimental vapour pressure (sublimation pressure) needs to be corrected to a sub-cooled liquid vapour pressure (see below) before comparison with estimated values; or use in atmospheric models. The use of Knudsen effusion linked to mass-spectrometry (KEMS) is well established for the measurement of partial pressures above alloy and intermetallic systems at high temperatures (Copland and Jacobson, 2001; Bencze et al., 2004); and has recently been used for the determination of vapour pressures of organic compounds at ambient temperatures (Booth et al., 2009). Within the atmospheric community several measurements have been made at ambient temperatures on mono- and di-carboxylic acids using a variety of techniques based upon the rate of evaporation of a compound under controlled conditions (Bilde et al., 2003; Koponen et al., 2007; Tao and McMurray, 1989; Cappa et al., 2007). However it is not clear at this time how these methods compare to more established techniques such as transpiration and Knudsen effusion.

1.2 Estimation of vapour pressures

Many methods for the estimation of vapour pressures have appeared in the literature and they have been periodically reviewed (Poling et al., 2001). However most of the methods are aimed at the requirements of the chemical industry and typically provide good results for volatile fluids, particularly if an experimental boiling point is available. The databases used to develop these methods are heavily biased towards mono-functional (or indeed non-functional i.e. hydrocarbon) compounds with relatively few examples of bi- or multi-functional compounds present. Hence the estimation methods tend to work best for compounds with one (or zero) functional groups and the relatively few compounds with two or more functional groups may well end up as outliers unless specific provision is made by the model to accommodate them (e.g. by including interactions between functional groups).

For some of the models, particularly the more complex group contribution methods, there is a significant danger of overfitting. This is a problem that is well understood by the chemometrics and cheminformatics community (Faber and Rajko, 2007); and occurs when a model starts to describe the noise in the data rather than the underlying trend in the data. An overfitted model will give very good results for the data used to fix the adjustable parameters but much worse results for new data outside the training set. The potential for overfitting depends upon the amount of training data, the number of adjustable parameters in the model and how well the mathematical form of the model relates to the true trend of the data. As group contributions models for vapour pressures (and other physical properties) become more complex then the ratio of independent data points to adjustable parameters is reduced. This results in a significant number of the parameters being fitted to small subsets of the data because the relevant structural feature is poorly represented in the overall data set. In these circumstances there is a danger that a statistically significant non-zero parameter value may be incorrectly obtained. To take a hypothetical example: a specific group contribution may be obtained by the fitting of vapour pressure data, for say 3 or 4 structurally related compounds (eg C4, C5, C6 dicarboxylic acids) all measured by the same researcher, using the same equipment over a period of one year. As experimentalists often report results for structurally related compounds together in a paper this scenario is quite possible but it does raise the question as to the definition of “independent” data. The fitted parameter may reflect biases in the work that would disappear if the data could be combined with that for a more structurally diverse range of dicarboxylic acids produced by a variety of researchers using different techniques. Hence the best way to avoid the danger of overfitting in group contribution models is to ensure that as many independent data points as possible are contributing to the fit for all the adjustable parameters.

Furthermore few, if any, of the vapour pressure estimation methods claim much accuracy below 100 Pa, and the errors increase significantly at lower pressures. Most of the estimation methods reported in the literature require a normal boiling point (T_b) which, for multifunctional compounds, will usually have to be estimated. Hence many estimation methods have two parts: estimation of T_b , followed by extrapolation from T_b down to the temperature of interest. For some of the compounds of atmospheric interest, the estimated T_b can be in excess of 700 K so a relatively small error in the slope of the line between T_b and 25°C can make a large difference to the predicted vapour pressures. Estimation methods of this type that have been used within the atmospheric science community include T_b estimation by the method of Nannoolal et al. (2004) (used in E-AIM: Extended Aerosol Inorganics Model, http://www.aim.env.uea.ac.uk/aim/ddbst/pcalc_main.php; see Wexler and Clegg, 2002); and the older group contribution method of Stein and Brown (1994), used in both E-AIM and EPI-Suite (EPA, 2009). This last method is a

modification of the method of Joback and Reid (1987) which has recently been used in the study of Camredon and Aumont (2006). E-AIM provides vapour pressure data either by using T_b by Stein and Brown (1994) combined with the vapour pressure equation of Myrdal and Yalkowsky (1997); or with T_b estimated by Nannoolal et al. (2004) coupled with either the vapour pressure equation of Nannoolal et al. (2008); or that of Moller et al. (2008), using the ARTIST software from the DDB. Camredon and Aumont (2006) reported use of the Myrdal and Yalkowsky (1997) equation while EPI-Suite uses a modified version of the Grain (1982) equation with a Fishtine factor (Lyman, 1985; Fishtine, 1963). Prediction of the absorptive partitioning of large-numbers of compounds requires automation of the methods for the estimation of vapour pressures (Aumont et al., 2005; Johnson et al., 2006). EPI-Suite can provide estimated boiling points and vapour pressures rapidly in batch mode using SMILES strings (EPA, 2009) as input. However the vapour pressure equation used by EPI-Suite has limited ability to predict the slope of a vapour pressure curve with respect to temperature due to the small number of values (6 in total) that the Fishtine factor can take. Also, apart from polyols, the application of the Fishtine factor to multifunctional compounds is undefined. In contrast the Nannoolal et al. (2008) method provides 130+ group contributions plus group interactions (all derived from experimental data) to predict the slope of the vapour pressure curve with temperature. To test whether this large number of fitted parameters (133 required for our test set of multifunctional compounds) clearly improved the vapour pressure predictions, a simplified form of the Nannoolal et al. (2008) equation in which all group interactions were dropped, and the number of structural groups required to describe the test set was reduced to 15, was created and tested against the other methods. This simplified equation was expected to produce inferior results to the full model although it does retain the functional form with temperature of the full method.

Some recently reported estimation methods provide vapour pressure values at low temperature without using a boiling point. In this latter category there are a number of methods that have been developed specifically for compounds of atmospheric interest, although in some cases the range of functionality is limited. These include the estimation method of Capouet and Muller (2006) which uses the estimated vapour pressure of a homologous hydrocarbon which is then corrected for the functionality present; and the group contribution method from Pankow and Asher (2008) specifically aimed at atmospherically important compounds.

The selection of a vapour pressure estimation method for use in the modelling of aerosol formation is always going to be a compromise between accuracy, complexity and coverage of all the required functional groups. Increasing the complexity of a model by adding in more adjustable parameters to improve accuracy can result in overfitting if taken too far. It is clear that there is no general agreement as to

the best vapour pressure method to use for compounds of atmospheric interest. Clegg et al. (2008) reviewed the predictions made by a wide range of vapour pressure estimation methods of a limited number of surrogate compounds of atmospheric interest and showed that the methods could give very divergent predictions for the same compound. The current work specifically aims to identify the estimation method likely to provide the most accurate vapour pressure predictions for the degradation products of atmospheric volatile organic compound (VOC) oxidation, thereby providing a recommendation for the method most suitable for vapour pressure prediction for the purposes of ambient gas-particle partitioning. The sensitivity of aerosol formation to variation in vapour pressure values will also be investigated. It is assumed that the organic compounds most likely to contribute to OA will be multifunctional and have low vapour pressures (<100 Pa) at ambient temperatures. This includes any components which may be present in OA by virtue of properties or processes other than their low volatility (e.g. reversible or irreversible reactive uptake or significant condensed phase reactions).

2 Methodology

Vapour pressure estimation methods based upon the principle of corresponding states seem unlikely to be able to give good predictions for complex multifunctional compounds of atmospheric importance. This is because these models require values for the compound's critical temperature (T_c) and pressure (P_c). These methods can give good results for volatile fluids, particularly if experimental T_c and P_c are available (Poling et al., 2001). However for the compounds of interest in this work critical values would have to be estimated, and the estimation methods (e.g. methods by Joback and Ambrose, Reid et al., 1987) typically use a T_b value; which will also be estimated. Hence there is considerable scope for an accumulation of errors. Also the database for experimental critical properties is much smaller than that for boiling points, and will contain even fewer multifunctional compounds. This suggests that any critical property estimation method will provide rapidly poorer estimates of T_c and P_c as the molecular functionality increases and the volatility decreases. Despite the logic of these arguments it is important to note that Camredon and Aumont (2006) obtained very acceptable results using the Lee-Kesler equation with T_c and P_c estimated by the method of Ambrose (Reid et al., 1987); although it is noticeable that the scatter of their data increases markedly at lower volatilities in agreement with the above argument. Despite our reservations about corresponding states methods we decided to test the same method against our test set of 45 multifunctional compounds to see how our results compared to those of Camredon and Aumont.

The bulk of this work will consider combined estimation methods where T_b is first estimated and then a vapour

pressure value at the required temperature is obtained by extrapolation from T_b using a vapour pressure equation. To assess the accuracy of the vapour pressure equations it is necessary to use low or very low vapour pressure data (below 100 Pa) to ensure a significant extrapolation from T_b . While vapour pressure data below 100 Pa are available for a range of multifunctional compounds, only a small number (e.g. some diols, polyethers, anisaldehyde and glycerol) also have an experimentally-determined T_b value. Unfortunately, it is therefore not possible to assess the accuracy of the vapour pressure equations independently of T_b values using a diverse set of multifunctional compounds. In this work, vapour pressure equations were first screened using low vapour pressure data for those multifunctional compounds for which experimental T_b values are available, supplemented by a more diverse set of additional compounds (mainly monofunctional); in combination referred to as Test Set 1. From this screening process, four independent equations were selected to be combined with three T_b estimation methods and evaluated against vapour pressure data for 45 multifunctional compounds; referred to as Test Set 2.

The boiling point estimation methods selected were all group contribution methods mentioned in Sect. 1.2 and span a wide range of complexity. The simplest method is that of Joback and Reid (1987) (JR method – 41 groups); the method of Stein and Brown (1994) (SB) is adapted from the JR method with additional groups (85 in total) and a correction for high boiling point values. The third method is that of Nannoolal et al. (2004) (N-Tb) which includes both primary and secondary groups along with group interactions (207 terms in total).

2.1 Selection of vapour pressure estimation methods

The number of vapour pressure equations reported in the literature that could be combined with estimated T_b values is large, although several equations are variations on each other. Two vapour pressure equations that have been widely quoted in the environmental literature are the Grain-Watson (GW) equation (Lyman, 1985; Grain, 1982) and a simplified version quoted by Baum (1998). The GW equation has the form:

$$\ln(p_i^0) = \frac{\Delta S_{\text{vap}}}{R} \left[1 - \frac{(3-2T_p)^m}{T_p} - 2m(3-2T_p)^{m-1} \cdot \ln(T_p) \right]. \quad (1)$$

Where $T_p = T/T_b$ and $m = 0.4133 - 0.2575 T_p$. In this section calculated vapour pressures (p_i^0) are in atmospheres, and T and T_b are the temperature and normal boiling point (both in K), respectively.

The simplified expression from Baum is:

$$\ln(p_i^0) = \frac{\Delta S_{\text{vap}}}{R} \left[1.8 \left(\frac{T_b}{T} - 1 \right) - 0.8 \ln \frac{T_b}{T} \right]. \quad (2)$$

Both expressions require a value for the latent entropy of vapourisation (ΔS_{vap}) at the normal boiling point. This can

be obtained from the latent enthalpy of vapourisation if it is known:

$$\Delta S_{\text{vap}} = \frac{\Delta H_{\text{vap}}}{T_b} \quad (3)$$

or can be estimated by a group contribution method: e.g. Joback and Reid (1987).

$$\Delta H_{\text{vap}} = 15.30 + \sum_i N_i H_i, \quad (4)$$

where $N_i H_i$ are the contributions for each group (i) and ΔH_{vap} is in kJ mol^{-1} (Poling et al., 2001).

Alternatively, various expressions based upon Trouton's Rule have been suggested e.g. Lyman (1985) proposed:

$$\Delta S_{\text{vap}} = K_f R \ln(82.06 \cdot T_b), \quad (5)$$

where K_f is a structural factor of Fishtine (1963).

The EPA software EPI-Suite uses the GW equation (Eq. 1) with a very similar expression to the above for ΔS_{vap} to calculate vapour pressure values. Unfortunately the Fishtine factor is only defined for a very limited range of functional groups and this form of the expression may not work well for complex multifunctional compounds. Vetere (1995) has proposed equations for ΔS_{vap} that are polynomials in T_b (Poling et al., 2001):

$$\Delta S_{\text{vap}} = A + B \log_{10}(T_b) + \frac{CT_b^{1.72}}{M}. \quad (6)$$

Where A , B and C take different values for hydrocarbons, alcohols/acids and all other polar compounds and M is the molecular weight (modified for halogen and phosphorus compounds).

The equation of Myrdal and Yalkowsky (1997) (MY) relates the latent entropy of vaporisation to the rigidity of the molecular structure and the number of hydrogen bonds:

$$\Delta S_{\text{vap}} = 86 + 0.4\tau + 1421 \times \text{HBN}. \quad (7)$$

Where τ is the effective number of torsional bonds and HBN is the hydrogen bond number. The vapour pressure equation takes account of heat capacity terms and has the form:

$$\log_{10}(p_i^0) = \frac{\Delta S_{\text{vap}}(T_b - T)}{19.1T} - \frac{[90.0 + 2.1\tau]}{19.1} \left(\frac{T_b - T}{T} - \ln \frac{T_b}{T} \right). \quad (8)$$

The most complex method considered here is that of Nannoolal et al. (2008) (will be referred to as the N-VP equation) which uses a group contribution method (same groups as in the N-Tb method mentioned above) to predict the slope of the vapour pressure curve:

$$\log_{10}(p_i^0) = (4.1012 + dB) \left[\frac{T_p - 1}{T_p - 0.125} \right]. \quad (9)$$

Where T_p is the reduced temperature ($=T/T_b$) and is referred to as T_{rb} in Nannoolal et al. (2008). The parameter dB , which

Table 1. The reduction of Nannoolal primary group parameters contributing to dB down to 15 values used to represent the compounds in Test Set 2 in the N-Sim estimation method.

Functional group	Nannoolal groups ^a	Range of C_i ^b	Mean contrib. ^b
Alcohol	33–36	0.7007–0.7584	0.7336
Aldehyde/Ketone	51, 52, 90, 92	0.1583–0.2558	0.2125
Ester	45–47	0.2928–0.3557	0.333
Ether	38 ^c	0.1085	0.185
Carbonate	79 ^d	0.3182	0.3182
Amine-aliphatic	40, 42, 43 ^e	–0.1027–0.2519	0.1143
Nitro	68, 69	0.2382–0.3758	0.3070
Chloride	25–29	0.0460–0.0937	0.0697
Hydrocarbon	1–18, 58–62, 88, 89 ^f	–0.0318–0.1121	0.0435

^a Other functional groups (Nannoolal group in brackets – see Nannoolal et al., 2008): phenol (37), carboxylic acid (44), amine-aromatic (41), nitrate (72), bromide-aromatic (31) and nitrile/cyanide (57) all transferred from the full method without modification.

^b All parameters rounded to 4 significant figures.

^c Excludes group 65 – an aromatic O in an aromatic ring as found in furan.

^d Excludes cyclic carbonates.

^e Excludes group 97 – secondary amine in a ring.

^f Hydrocarbon group includes all groups in the section headed carbon except those for triple bonds and cumulative double bonds.

adjusts the slope of the vapour pressure curve, is estimated from the structure of the molecule using group contributions:

$$dB = \left(\sum N_i C_i + GI \right) - 0.176055, \quad (10)$$

where the first term in the brackets is the sum of group contributions for both primary and secondary groups and the second term refers to a group interaction contribution:

$$GI = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^m \frac{C_{i-j}}{m-1}, \quad (11)$$

where $C_{i-j} = C_{j-i}$; and m , n are the total number of interacting groups and the number of (non-hydrogen) atoms in the molecule, respectively.

In summary the seven vapour pressure equations that were assessed in the first screening were:

- The GW equation (Eq. 1) with ΔS_{vap} given by the formalism of Vetere (Eq. 6).
- The equation of Baum (Eq. 2) with ΔS_{vap} given by the formalism of Vetere (Eq. 6).
- The GW equation (Eq. 1) with $\Delta S_{\text{vap}} = K_f R \ln(82.06 \times T_b)$ (Eq. 5).
- The equation of Baum (Eq. 2) with $\Delta S_{\text{vap}} = K_f R \ln(82.06 \times T_b)$ (Eq. 5).
- The GW equation (Eq. 1) with ΔS_{vap} given by a group contribution method for ΔH_{vap} (Eq. 4).
- The MY equation (Eqs. 7, 8).
- The N-VP equation (Eqs. 9–11).

This work will also consider some variants on the Nannoolal method; a corresponding states method (Lee-Kesler), and two methods that do not require a boiling point. The accuracy of these methods will be compared to the accuracy of the best combined methods once these have been found using Test Set 2.

The full N-VP model has primary and second order group contributions and group interactions (GI in Eqs. 10 and 11); some 133 parameters are required to describe the compounds of Test set 2. To test whether all these parameters are important for the estimation of the vapour pressures of the 45 multifunctional compounds a simplified vapour pressure equation was created with 15 parameters (see Table 1). This simplified Nannoolal model (referred to in this work as the N-Sim method) was designed to be a version without second order group contributions or group interactions and with a minimal number of primary group contributions. The aim was to find out if these addition parameters, which might be very important for determining the vapour pressures of the relatively volatile hydrocarbons and monofunctional compounds which will dominate the experimental data used in the parameter fitting, had a detectable impact upon the predictions for the low volatility multifunctional compounds of primary interest to this work. The primary groups are required to describe the molecule and it is important that all heavy atoms (C, O, N, Cl) are accounted for. However the use of 31 groups to describe hydrocarbon structures is probably driven by detailed vapour pressure data for these compounds and we decided to simplify the method by taking an average value (see Table 1). We retained all the main non-hydrocarbon functional groups (with some averaging, e.g. for alcohols) to account for all heavy atoms and because the size

of the group contributions for the functional groups were so much larger than those for the hydrocarbon groups (see Table 1). The second order groups are corrections for specific structural features and generally give small group contribution values. Although the group interactions could potentially be important for our test set and some of the group contributions are quite large it should be noted that their contribution to the slope of the vapour pressure curve is inversely proportional to the number of heavy atoms in the molecule. Hence their impact upon the vapour pressures of the relatively large molecules in Test set 2 may be limited. On this basis the second order group contributions and group interactions were dropped from the N-Sim method.

The method of Moller et al. (2008) (will be referred to as the M-VP method) is an improvement of the Nannoolal et al. (2008) (N-VP) method. It features an additional term to improve predictions for aliphatic alcohols and carboxylic acids; new size dependent groups to improve predictions for several functional groups, and new hydrocarbon groups. Rewriting Eq. 9 and adding the extra term gives;

$$\log_{10}(P_i^0) = B' \frac{T - T_b}{T - C(T_b)} + D' \ln\left(\frac{T}{T_b}\right) \quad (12)$$

where the second term on the right side is the new term for carboxylic acids and alcohols and D' is set to zero when they are not present. In the N-VP method, $C(T_b) = T_b/8$ (see Eq. 9), but this is replaced with the following term:

$$C(T_b) = -2.65 + \frac{T_b^{1.485}}{135} \quad (13)$$

and B' and D' are obtained by the summation of the appropriated group contributions, including size dependant groups and group interactions. All group contributions were refitted to the above equations.

The Lee-Kesler method (or LK method – Reid et al., 1987 – Sect. 7.2) requires the critical temperature (T_c) and critical pressure (P_c) for each compound, and these properties were calculated by the estimation method of Ambrose (Reid et al. (1987) Sect. 2.2) using the best available estimated T_b values (provided by the N-Tb method – see below). The vapour pressure equation is:

$$\ln P_{ir}^0 = f^{(0)}(T_r) + \omega f^{(1)}(T_r) \quad (14)$$

where P_{ir}^0 is the reduced vapour pressure ($=P/P_c$ with P and P_c both in atmospheres) and T_r is the reduced temperature ($=T/T_c$ with T and T_c both in K). ω is Pitzer's acentric factor, while $f^{(0)}$ and $f^{(1)}$ are polynomials in T_r :

$$f^{(0)} = 5.92714 - \frac{6.09648}{T_r} - 1.28862 \ln T_r \quad (15)$$

$$+ 0.169347 T_r^6$$

$$f^{(1)} = 15.2518 - \frac{15.6875}{T_r} - 13.4721 \ln T_r \quad (16)$$

$$+ 0.43577 T_r^6$$

and ω is estimated from the T_b and critical properties using the following three equations (Reid et al., 1987 – Sect. 2.3):

$$\omega = \frac{\alpha}{\beta} \quad (17)$$

$$\alpha = -\ln P_c - 5.92714 + \frac{6.09648}{\theta} + 1.28862 \ln \theta - 0.169347 \theta^6 \quad (18)$$

$$\beta = 15.2518 - \frac{15.6875}{\theta} - 13.4721 \ln \theta + 0.43577 \theta^6 \quad (19)$$

where $\theta = T_b/T_c$.

The two methods for the estimation of vapour pressures that don't require a boiling point value are SIMPOL.1 (Pankow and Asher, 2008) and the method of Capouet and Muller (2006). The SIMPOL.1 method is a group contribution method where each group has a specified temperature dependence:

$$\log_{10}(P_i^0) = b_0(T) + \sum v_{k,i} b_k(T). \quad (20)$$

Where k takes the values 0, 1, 2, ... up to 30 and the term with $k=0$ is the zeroeth group or constant term. The non-zero k values refer to structural features in the molecule (aromatic ring, non-aromatic ring, aldehyde, ketone etc.) and $v_{k,i}$ is the number of times each structural feature occurs. Each $b_k(T)$ is a polynomial in temperature (T):

$$b_k(T) = \frac{B_{1,k}}{T} + B_{2,k} + B_{3,k} T + B_{4,k} \ln(T). \quad (21)$$

The estimation method of Capouet and Muller (2006) (CM method) uses the vapour pressure of a homologous hydrocarbon (estimated if required) which is then corrected for the functionality present:

$$\log_{10}(P_i^0) = \log_{10}(P_{hc}^0) + \sum_{k=1}^n v_{k,i} \tau_k(T). \quad (22)$$

Where the first term on the right refers to the vapour pressure of the hydrocarbon in atmospheres and the second term is the correction due to the functional groups. Unfortunately the authors only provide $\tau_k(T)$ for a limited range of functional groups (carbonyl, nitrate, hydroperoxy, hydroxyl, carboxylic acid and PAN) so this CM method could only be used on a small number of multifunctional compounds. SIMPOL.1 covers a wider range of functionality and a better comparison could be made with the other estimation methods.

2.2 Selection of vapour pressure data

Several sets of vapour pressure data have been collected by the atmospheric community and used to develop estimation methods (Pankow and Asher, 2008; Asher and Pankow, 2006; Camredon and Aumont, 2006; Capouet and Muller, 2006). However these sets often contain data from a mixture

of sources, including secondary sources that provide vapour pressure correlations rather than original experimental data. Secondary sources typically provide little or no information about the primary data upon which the correlations are based. In the current work two sets of vapour pressure data were collected solely from the primary literature.

Test Set 1 was used to choose the best vapour pressure equations from the seven (A to G) listed in Sect. 2.1, and the selected compounds had to meet the following criteria:

- (1) Experimental vapour pressure data at pressures below 40 Pa are available. In most cases several points collected as part of a set were required. A few values above 40 Pa have been included.
- (2) An experimental melting point is available; either from the primary literature, or from the Detherm database (www.dechema.de/en/detherm.html), and demonstrates that the measurements in (1) were made on a liquid.
- (3) An experimental boiling point is available and in most cases is supported by at least one set of experimental vapour pressure data approaching atmospheric pressure; or at least close to atmospheric pressure ($>9 \times 10^4$ Pa) so that the boiling point can be confidently obtained by extrapolation.

Members of this test set were selected with the aim of maximising the number of functional groups represented. In general no more than two examples were used for each functional group, although multifunctionality was counted separately (so there are two alcohols, two diols and a triol in this test set). Very few multifunctional compounds have both well-established boiling points and experimentally determined very low-vapour pressure data so most members of this test set were mono-functional; two hydrocarbons were included for completeness.

These criteria inevitably excluded a wide range of compounds from this test set. In addition to the exclusion of most multifunctional compounds the selected molecules tend to have a low molecular weight and the criteria that they need to be stable at the experimental boiling points may exclude some functional groups (e.g. nitrate, some nitro compounds and some aldehydes) which encourage decomposition at elevated temperatures. The requirement of liquid phase vapour pressure data below 40 Pa means that very regular molecules are excluded because they will have relatively high melting points (an extreme example is camphor). In the case of aromatics, alkyl substituents disrupt the crystal packing and lower the melting point so several of these appear in the list but more regular aromatics (e.g. benzoic acid) are excluded because of this requirement.

The compounds of Test Set 1 are listed in Table 2 along with their vapour pressure data, normal boiling point and melting point. Data sources are provided in the supplementary material: <http://www.atmos-chem-phys.net/10/749/2010/acp-10-749-2010-supplement.pdf>.

Compounds in Test Set 2, shown in Table 3, were used to test the accuracy of the combined (T_b and vapour pressure equation) estimation methods when applied to multifunctional compounds. In selecting experimental vapour pressure data for this test set the following criteria were followed:

- (1) The compounds must contain two or more non-hydrocarbon functional groups.
- (2) The vapour pressure data should consist of two or more consistent points below 100 Pa.
- (3) The paper describing the measurement of the vapour pressures should make it clear that the measurements were done on a liquid. If there is any ambiguity or if it was clear that the measurements were made on a solid then an experimental melting point from a primary source is required.
- (4) If the compound is a solid at the temperature of measurement then either the melting point should be within 30 K of the measurement temperature or a reliable experimental value for the latent enthalpy (ΔH_{fus}) or entropy (ΔS_{fus}) of fusion at the melting point should be available.

Using these criteria, data for 47 compounds were found in the primary literature. Two compounds were eventually dropped because estimated predictions for these compounds were closely correlated to those for several other compounds. The 45 remaining compounds with their melting points and vapour pressure data (temperature and pressure ranges after correction) are listed in Table 3. Data sources for Table 3 are provided in the Supplementary material: <http://www.atmos-chem-phys.net/10/749/2010/acp-10-749-2010-supplement.pdf>. In a few cases the authors represent their data as a correlation, rather than quoting the measured raw data (e.g. Lei et al., 1999). In these cases each dataset is usually represented by two points at the extremes of the experimental temperature range. For those compounds where the vapour pressures quoted are that of the solid (see Table 3) the corresponding sub-cooled liquid vapour pressures (SCL-VP) were obtained by applying the correction of Prausnitz et al. (1986):

$$\ln(p_i^0) = \ln(p_{\text{SCL}}) = \ln(p_{\text{S}}) - \frac{\Delta S_{\text{fus}}}{R}(1 - T_{\text{m}}/T) - \frac{\Delta C_p}{RT}(T_{\text{m}} - T) + \frac{\Delta C_p}{R} \ln\left(\frac{T_{\text{m}}}{T}\right). \quad (23)$$

Where p_{SCL} and p_{S} are the SCL-VP and experimental solid (sublimation) vapour pressure (in atmospheres), respectively; ΔS_{fus} is the entropy of fusion; and ΔC_p is the best estimate of the underlying change in heat capacity between the liquid and solid state at the melting point, and T_{m} is the melting point temperature (used instead of the triple point temperature).

Table 2. The compounds of Test Set 1^a and their properties.

Compound name	T_b (K)	T_m (K)	Exp. T range (K)	Exp. p range (Pa)
1,1'-Oxobisbenzene	531.2	300.4	309.15–329.55	5.33–33.46
1,1'-Oxobisethane	307.58	156.85	163.15–187.95	1.08–44
1,2-Ethanediol	470.79	260.6	263.59–313	0.474–41.9
1,2,3-Propanetriol	563	291.05	291.18–319.19	0.0095–0.228
1,2-Dichlorobenzene	453.63	256.4	256.51–283.47	7.06–67.1
1,6-Hexanediol	525.95	314.6	327–365	2.31–51.5
1-Methyl-2-nitrobenzene	495.3	269	274–303.5	2.371–29.66
2-Butanone	352.74	186.5	208.83–215.51	18.62–36.29
2-Ethyl phenol	477.67	269.8	278.13–302.68	3.35–30.8
2-Octanone	446.2	229.85	243.15–298.15	1.5–187
3-Methyl phenol	475.42	284	284.15–306.4	5.87–36.8
Benzyl alcohol	478.6	257.6	282.9–308.15	3.06–28
Cyclohexyl formate	435.55	201	248.25–268.44	5.59–40.7
Heptanoic acid	496.15	266	270.4–328.2	0.1–28.29
Hexanoic acid	479	269.15	271.4–313.2	0.4–25.73
Limonene	450.8	177.1	243.5–273.2	1.91–31.02
Linalool	471.35	?? ^b	273.35–303.14	2.49–42.2
<i>n</i> -Decanal	489	268	282.8–307.2	4.83–35.41
<i>n</i> -Octanenitrile	478.4	227.6	283.2–298.2	11.92–38.98
Phenethyl alcohol	492.05	246.2	288.15–318.15	2.8–40
Propylbenzene	432.39	173.6	223.15–263.15	0.61–36.2

^a For data sources see the Supplementary material (<http://www.atmos-chem-phys.net/10/749/2010/acp-10-749-2010-supplement.pdf>).

^b T_m unknown but low (below 273 K).

This correction has been discussed at some length by Capouet and Muller (2006) who pointed out that if T_m is close to the experimental temperature (T) then the last two terms in Eq. (23) tend to cancel. In this work, if T was within 30 K of T_m then it was considered that the last two terms could be ignored and an estimated ΔS_{fus} (using the method of Myrdal and Yalkowsky, 1997) could be used; though an experimental value was preferred if available. For those compounds where $T_m - T$ was greater than 30 K, an experimentally derived ΔS_{fus} was required:

$$\Delta S_{\text{fus}} = \frac{\Delta H_{\text{fus}}}{T_m}, \quad (24)$$

where ΔH_{fus} is the enthalpy of fusion and is usually obtained by differential scanning calorimetry. The heat capacity terms were also included with ΔC_p obtained either by a simple linear extrapolation of experimental heat capacity data for both the solid and liquid phase up to the melting point; or estimated from liquid heat capacity values obtained using the group contribution method of Ruzicka and Domalski (1993) and solid heat capacity data from the power law of Goodman et al. (2004). Details for all compounds requiring these corrections are provided in the Supplementary material: <http://www.atmos-chem-phys.net/10/749/2010/acp-10-749-2010-supplement.pdf>. In general the heat capacity correction was very small compared with the

entropic term but for phloroglucinol and 2-hydroxybenzoic acid it is quite substantial; in part because of the high melting points of these substances.

A basic error analysis was done for several compounds to estimate the impact of potential errors in the key parameters of Eq. 23. Assuming potential errors of $\pm 40\%$ for estimated quantities (ΔC_p and some ΔS_{fus} values – see Table 3); $\pm 20\%$ for experimental ΔS_{fus} and ± 2 K for T_m then for most compounds the calculated error in the final sub-cooled vapour pressure values was within $\pm 50\%$; although for a couple of compounds (Phloroglucinol and 2-Hydroxybenzoic acid) the potential error was substantially higher (up to a factor of three). For most compounds these potential errors are of the same order as the experimental scatter in low vapour pressure values, e.g. see Booth et al. (2009), and are small compared to the differences between the estimated and experimental vapour pressures seen in this work; but for Phloroglucinol and 2-Hydroxybenzoic acid the errors may be more significant and could influence the results for these compounds.

Table 3. Vapour pressure data and physical properties of the 45 compounds of Test Set 2.

No.	Compound name	CAS code	T_m (K)	T range (K) ^a	p range (Pa) ^a	n^f	Method ^g	ΔS_{fus}^h	Final p range (Pa) ⁱ
1	1,2-Pentanediol	5343-92-0	?? ^c	289.2–336.2	1.35–98.44	13	TR		1.35–98.44
2	1,2,3-Trihydroxypropane	56-81-5	291.8	298.75–341.35	0.0249–1.97	14	KE		0.0249–1.97
3	1,4-Butanediol	110-63-4	?? ^c	329.2–351.2	13.85–84.55	9	TR		13.85–84.55
4	2-(Methylamino)ethanol	109-83-1	268.6	274.9–296.3	13.94–92.12	9	TR		13.94–92.12
5	2,4-Pentanediol	625-69-4	?? ^c	297.2–330.5	4.8–87.35	12	TR		4.8–87.35
6	2,6-Dinitrotoluene	606-20-2	329.55	277.15–323.15	0.00342–1.718	6	TR	58.50 ^c	0.0127–1.974
7	2-Aminoethanol ^d	141-43-5	283.8	279.0–306.3	8.48–83.66	7	TR		8.48–83.66
8	2-Aminonitrobenzene	88-74-4	342.5	313.5–342.3	0.71–12.33	7	TR	47.04 ^b	1.20–12.37
9	2-Chloro-3,5-dimethoxy-4-HBA ^b	76341-69-0	469.65	293.15–323.15	0.046–0.9	2	GC-RT	– ^d	0.046–0.9
10	2-Chloropropionic acid	598-78-7	266.2	287.4–308.4	13.36–82.72	8	TR		13.36–82.72
11	2-Hydroxybenzoic acid	69-72-7	431.8	307.05–323.71	0.0682–0.468	7	KE	56.97 ^c	0.443–2.434
12	2-Phenylbromide-TEGME ^b	929259-37-0	?? ^c	333.2–369.9	0.061–1.446	19	TR		0.061–1.446
13	3-(2-Methoxyphenyl)PA ^b	6342-77-4	360.46	331.156–347.165	0.156–1.096	9	KE	70.27 ^b	0.329–1.514
14	3-(3,4-Dimethoxyphenyl)PA ^b	2107-70-2	370.85	352.178–366.163	0.0664–0.4115	20	KE	87.31 ^b	0.1159–0.4707
15	3,4-Dichloronitrobenzene	99-54-7	314.1	316.3–346.5	9.1–74.69	11	TR		9.1–74.69
16	3,4-Dihydroxychlorobenzene	2138-22-9	364.15	293.15–323.15	0.6–8.7	2	GC-RT	– ^d	0.6–8.7
17	3,5-di-tert-Butylcatechol	1020-31-1	372.8	313.2–346.2	0.0732–3.17	11	TR	64.65 ^c	0.270–5.58
18	3,7-Dimethyl-7-hydroxyoctanal	107-75-5	?? ^e	283.355–332.65	0.117–15.87	5	KE		0.117–15.87
19	3-Chloro-2,6-dimethoxyphenol	18113-22-9	308.65	293.15–323.15	0.52–7.1	2	GC-RT	– ^d	0.52–7.1
20	3-Chloroaminobenzene	108-42-9	262.8	291.2–325.3	6.02–76.82	10	TR		6.02–76.82
21	3-Hydroxypropanenitrile	109-78-4	180.4	306.3–331.4	15.64–91.47	13	TR		15.64–91.47
22	3-Nitro-3-(4-nitrophenyl)-pentane	204189-06-0	?? ^c	321.4–358.1	0.063–1.920	8	TR		0.063–1.920
23	3-Nitrobenzoic acid	121-92-6	413	347.16–361.16	0.215–0.905	9	KE	51.82 ^c	0.604–2.03
24	3-Nitrophenol	554-84-7	370	357.2–369.3	12.05–35.13	7	KE	46.76 ^b	14.74–35.51
25	4-Aminobenzoic acid	150-13-0	461.4	359.14–382.56	0.1–1.0	2	KE	45.30 ^c	0.28–2.31
26	Anisaldehyde	123-11-5	272.25	283.95–322.95	1.32–30.4	11	KE		1.32–30.4
27	Benzyl salicylate	118-58-1	295.15	295.45–342.95	0.016–1.24	11	KE		0.604–1.24
28	Dibutyl phthalate	84-74-2	238.15	293.05–317.15	0.00191–0.0335	9	KE		0.00191–0.0335
29	Ethyl vanillin	121-32-4	351.2	323.35–337.45	0.47–2.75	6	KE	72.45 ^b	0.96–3.78
30	Eugenol	97-53-0	262.8	285.45–326.75	0.64–20.00	16	KE		0.64–20.00
31	Glutaric acid	110-94-1	371	348.15–363.15	0.224–1.19	8	KE	62.0 ^b	0.366–1.397
32	Glycerine carbonate	981-40-8	?? ^c	330.2–398.5	0.29–46.94	20	TR		0.29–46.94
33	Heliotropin	120-57-0	310.2	293.45–326.85	0.39–11.60	10	KE	56.5 ^a	0.58–11.60
34	Isoamyl salicylate	87-20-7	?? ^e	287.95–328.55	0.22–8.93	20	KE		0.22–8.93
35	Ketol	28746-99-8	303.13	308.22–330.4	1.23–6.64	7	KE		1.23–6.64
36	Methyl anthranilate	134-20-3	298.65	299.45–319.15	2.32–13.73	16	KE		2.32–13.73
37	Musk ambrette	83-66-9	358.15	328.55–345.45	0.141–0.973	5	KE	70.3 ^a	0.303–1.33
38	N-methyl-diethanolamine	105-59-9	252.2	293.69–353.0	0.61–80.9	15	ST		0.61–80.9
39	p-Acetylanisole	100-06-1	311.65	313.55–333.45	5.60–25.5	13	KE		5.60–25.5
40	Phloroglucinol	108-73-6	491.8	381.31–404.58	0.1–1.0	10	KE	70.15 ^c	0.82–5.06
41	Pimelic acid	111-16-1	377.5	358.15–371.66	0.126–0.675	8	KE	80.26 ^b	0.212–0.786
42	Pinonaldehyde	2704-78-1	?? ^c	263.15–278.15	0.09–0.6	4	KE		0.09–0.6
43	Tetraethylene glycol	112-60-7	263.8	323.23–398.23	0.173–44.7	12	TR		0.173–44.7
44	Triacetin	102-76-1	276.4	284.2–318.2	0.0512–2.08	8	PM		0.0512–2.081
45	Triethylene glycol dinitrate	111-22-8	?? ^c	303.4–348.0	0.025–2.21	9	PM		0.025–2.21

^a Data as reported in the literature – see supplementary material (<http://www.atmos-chem-phys.net/10/749/2010/acp-10-749-2010-supplement.pdf>) for data sources.

^b Full names: compounds No. 9=2-chloro-3,5-dimethoxy-4-hydroxybenzaldehyde; No. 12=(2-phenylbromide)-triethyleneglycol-monomethylether; No. 13=3-(2-methoxy phenyl)propionic acid; No. 14=(3,4-dimethoxyphenyl)propionic acid.

^c Melting point unknown, but from the vapour pressure source it is clear that the measurements were made on a liquid.

^d Authors note that some vapour pressure measurements were made on a sub-cooled liquid aided by very slow crystallization (Kapteina et al., 2005).

^e Melting point unknown but there is circumstantial evidence that the measurements were made on a liquid – see supplementary material for details (<http://www.atmos-chem-phys.net/10/749/2010/acp-10-749-2010-supplement.pdf>).

^f Number of datapoints.

^g Experimental method: KE=Knudsen effusion; TR=transpiration; GC-RT=gas chromatography-retention time; ST=static method; PM=piston manometer.

^h ΔS_{fus} at T_m in J/molK; correction of sublimation pressures to SCL-VP values: (a) correction ignoring the last two terms in Eq. (23) and using an estimated ΔS_{fus} ; (b) correction ignoring the last two terms in Eq. (23) and using an experimental ΔS_{fus} ; (c) correction using the full Eq. (23) with an experimental ΔS_{fus} , see supplementary material (<http://www.atmos-chem-phys.net/10/749/2010/acp-10-749-2010-supplement.pdf>) for details; (d) no correction required as experimental method provides SCL-VP values directly.

ⁱ Final pressure range after any corrections.

Table 4. Results (ΔT_b values in K) for the screening of seven vapour pressure models against Test Set 1.

Compound name	Vapour pressure equation ^a						
	A	B	C	D	E	F	G
1,1'-Oxobisbenzene	13.1	7.6	6.2	0.8	13.1	5.5	7.1
1,1'-Oxobisethane	4.9	1.3	6.3	2.4	3.7	-7.5	2.8
1,2-Ethanediol	-10.6	-13.9	-4.4	-8.3	-1.6	-11.9	4.8
1,2,3-Propanetriol	-10.3	-16.7	-3.1	-10.6	-21	4.4	-11.4
1,2-Dichlorobenzene	-2.2	-6.5	-6.2	-10.5	-3.0	-7.8	1.6
1,6 Hexanediol	7.1	3.5	15.9	11.5	21.9	26.2	20.3
1-Methyl-2-nitrobenzene	5.9	2.6	1.0	-4.3	-37.0	0.2	1.6
2-Butanone	1.9	-1.4	0.0	-3.4	3.1	-4.5	0.9
2-Ethyl phenol	-20.4	-23.6	2.3	-2.2	2.0	-1.0	8.3
2-Octanone	8.9	4.2	3.3	-1.4	11.6	-5.8	-0.4
3-Methyl phenol	-17.0	-20.0	5.9	1.7	19.5	1.5	12.7
Benzyl alcohol	-13.6	-16.7	-5.4	-9.1	69.3	5.7	9.0
Cyclohexyl formate	-3.0	-7.0	-0.3	-4.7	6.8	-3.8	-1.6
Heptanoic acid	-3.4	-7.8	46.7	38.8	97	18.2	9.9
Hexanoic acid	-4.3	-8.5	41.8	34.8	84.2	14.1	8.7
Limonene	5.8	0.6	3.2	-2.2	15.0	-6.5	3.9
Linalool	-16.6	-19.9	-11.1	-14.9	-17.4	1.0	0.6
<i>n</i> -Decanal	15.5	10.5	8.7	3.8	19.8	-1.8	0.9
<i>n</i> -Octanenitrile	7.0	2.5	3.3	-1.3	8.0	-3.8	0.6
Phenethyl alcohol	-18.1	-21.3	-9.6	-13.5	-1.6	2.9	2.0
Propylbenzene	4.4	-0.9	-4.2	-9.3	9.2	-9.9	-1.6
Bias	-2.13	-6.25	4.78	-0.09	14.41	0.73	3.84
Standard Dev.	10.88	10.42	14.6	13.79	32.46	9.44	6.45

^a Models tested (see text): (A) the GW equation with ΔS_{vap} given by Vetere's equations; (B) the Baum equation with ΔS_{vap} given by Vetere's equations; (C) the GW equation with $\Delta S_{\text{vap}}=K_f R \ln(82.06 \cdot T_b)$; (D) the Baum equation with $\Delta S_{\text{vap}}=K_f R \ln(82.06 \cdot T_b)$; (E) the GW equation with ΔS_{vap} given by the Joback group contribution method; (F) the MY method; (G) the N-VP method.

3 Results

3.1 Evaluation of vapour pressure equations against Test Set 1

For each compound in Test set 1 the predicted vapour pressure was calculated for each point in the dataset using the experimental T_b value. The T_b value was then changed to the value giving the closest fit of estimated vapour pressure (p_{est}) to experimental vapour pressure (p_{exp}), as defined by minimising the objective function – Eq. 25. The experimental T_b was then subtracted from the fitted T_b (to give ΔT_b) and the resulting values analysed statistically for each method across all the compounds (see Table 4).

$$\text{OF} = 1 - \frac{1}{n} \sum_{i=1}^n \left(\frac{p_{\text{est}}}{p_{\text{exp}}} \right), \quad (25)$$

where the summation is over the n experimental points in the dataset for the specific compound being considered.

The difference between the boiling point value required to fit the experimental vapour pressure and the true (experimental) boiling point is a measure of the accuracy of extrapolation from the boiling point to the experimental temperatures and is recorded in Table 4 under ΔT_b for all seven methods.

Approach G (see Sect. 2.1), using the N-VP equation (Eqs. 9–11), gives the lowest standard deviation, while the smallest bias is provided by method D (vapour pressure calculated using the Baum equation – Eq. 2 – with ΔS_{vap} from Eq. 5). The best three methods with regard to bias are D, F and A; while the best three methods as judged by scatter (standard deviation) are G, F and B. This provided five methods that were better than the remaining two based on either bias or scatter. However the ΔT_b values of those methods using a common expression for ΔS_{vap} with the GW equation (Eq. 1) or the Baum equation (Eq. 2) were very closely correlated. ΔT_b values calculated using methods A and B showed correlation coefficients ≥ 0.99 , as did values from methods C and D. There is no benefit in testing both Eqs. (1) and (2) against the multifunctional compounds. As the methods using Eq. (2) gave smaller standard deviations than the corresponding methods using Eq. (1) the Baum equation (Eq. 2) was selected for further comparison and methods B, D, F and G were chosen to be evaluated against Test Set 2.

3.2 Evaluation against data for multifunctional compounds (Test Set 2)

The four estimation methods above were selected for their ability to extrapolate from normal boiling point down to

around ambient temperatures. This section describes the identification of the preferred combined method of estimating vapour pressures for a test set of multifunctional compounds (Test Set 2).

Normal boiling point (T_b) values were calculated by three methods for all 45 compounds. Values calculated by the N-Tb method were verified against the E-Aim website; those calculated using the SB method were verified against EPI-Suite and the JR values were verified against calculated data from Chem-Draw Ultra version 10.0 (Cambridgesoft, 2005). As the published descriptions of the latter two methods do not include all the groups required to cover the structural features for the 45 compounds these checks were useful for ensuring that the more complex functional groups (such as carbonate) were represented in a way that was consistent with these readily available implementations. Each set of T_b values were used as input for the prediction of vapour pressures at the experimental temperatures for each dataset. The accuracy of the prediction for each compound was summarised by calculating a mean bias error (MBE; see Camredon and Aumont, 2006, Eq. 26).

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^n [\log_{10} p_{\text{est}} - \log_{10} p_{\text{exp}}] \quad (26)$$

Where the summation is over all the datapoints for that compound. Hence a single parameter is calculated for each dataset independent of how many datapoints are in the dataset. The MBE values are summarised for all 12 combined methods in Table 5 along with the calculated means and standard deviation.

The different estimation methods are compared in the scatter plots shown in Figure 1 (regression coefficients are tabulated in Table 6) and in Fig. 2 where the data in Table 5 are summarised in box-whisker plots. In this box-whisker plot (and in Figs. 5 and 6) the central line of the box marks the median value with the upper and lower bounds of the box marking the upper and lower quartiles, respectively. The notch indicates the 95% confidence limit for the median value. Hence if the notches of two estimation methods don't overlap then the median values are different at the 95% confidence level. In this plot (and in Fig. 1) it is clear that the four methods that included the estimation of T_b using the JR group contribution method have a significantly different distribution from the other methods. The JR method is consistently over-estimating T_b leading to a substantial underprediction of vapour pressure. The most meaningful comparison is with the SB method which was derived from the JR method but with a correction for high temperature boiling points. This work shows that the correction introduced by Stein and Brown is definitely required for these low volatility multifunctional compounds. In contrast to this general trend it is clearly seen in Fig. 1 that the JR method gives unreasonably high vapour pressures for one compound in Test set 2: compound 32, Glycerine carbonate. This reflects the difficulty in

handling groups such as carbonate in group contribution estimation methods. In the JR method this group is considered as a combination of an ester and an ether group, while in the EPI-Suite implementation of the SB method it is considered as two ester groups. Hence the JR method underestimates the T_b value (giving the high estimated vapour pressures clearly seen in all four panels of Fig. 1), while the SB methods overestimates the T_b giving vapour pressures that are too low (see Table 5 for the MBE values). The N-Tb method has a specific group contribution for cyclic carbonates obtained from experimental data and gives a better estimate of the experimental vapour pressures (compared to T_b by the SB method) when used with 3 out of the 4 vapour pressure equations.

It is clear from Figs. 1 and 2 (and Tables 5 and 6) that the combined method using the N-VP equation with T_b by N-Tb gives the most accurate values; based upon minimum standard deviation in Table 5 and best R^2 value (for the full set of 45 compounds) in Table 6. The MY vapour pressure equation is the second best out of the four vapour pressure equations; but this method does have a bias towards overestimating vapour pressures (note the relatively large B coefficient in Table 6 for the N-Tb/MY method compared to the N-Tb/N-VP method).

Figure 3 displays the results for those methods that did not require a T_b value: SIMPOL.1 and the CM method. The former method could be used on 36 of the 45 compounds (mainly resulting from a lack of group contributions for halogen, carbonate and nitrile) and gave results much more scattered than, either the N-VP equation or the M-VP equation with T_b by N-Tb (see R^2 values in Table 6).

The CM method could only be used on 9 compounds (containing alcohol, carbonyl or acid groups) out of 45 and gave more scatter (as judged by the R^2 values in Table 6) than the N-Tb/M-VP combined method and very similar scatter to the N-Tb/N-VP method for this limited subset of compounds.

Figure 4 compares selected vapour pressure equations using the best T_b estimation method identified from the above analysis of the performance of the combined methods on Test set 2 (the N-Tb method). Despite the reservations about the suitability of vapour pressure methods requiring critical properties for low volatility multifunctional compounds outlined above (see Sect. 2), the LKA method gave quite reasonable results although the vapour pressure equation clearly contributes more to the scatter of the data than the N-VP or M-VP equations (leading to a reduction in R^2 from more than 0.79 to about 0.74 – see Table 6). Figure 4, panel b shows that the N-Tb/LKA model data does become more scattered at lower experimental vapour pressures but the effect is less marked than that seen by Camredon and Aumont (2006) (see their Fig. 4, panel a).

The other conclusion that may be drawn from Fig. 4 is that for Test set 2 there is little difference in the accuracy of the N-VP, M-VP and N-Sim vapour pressure equations when used with the N-Tb estimation method. This strongly suggests that the differences between the estimated and experimental

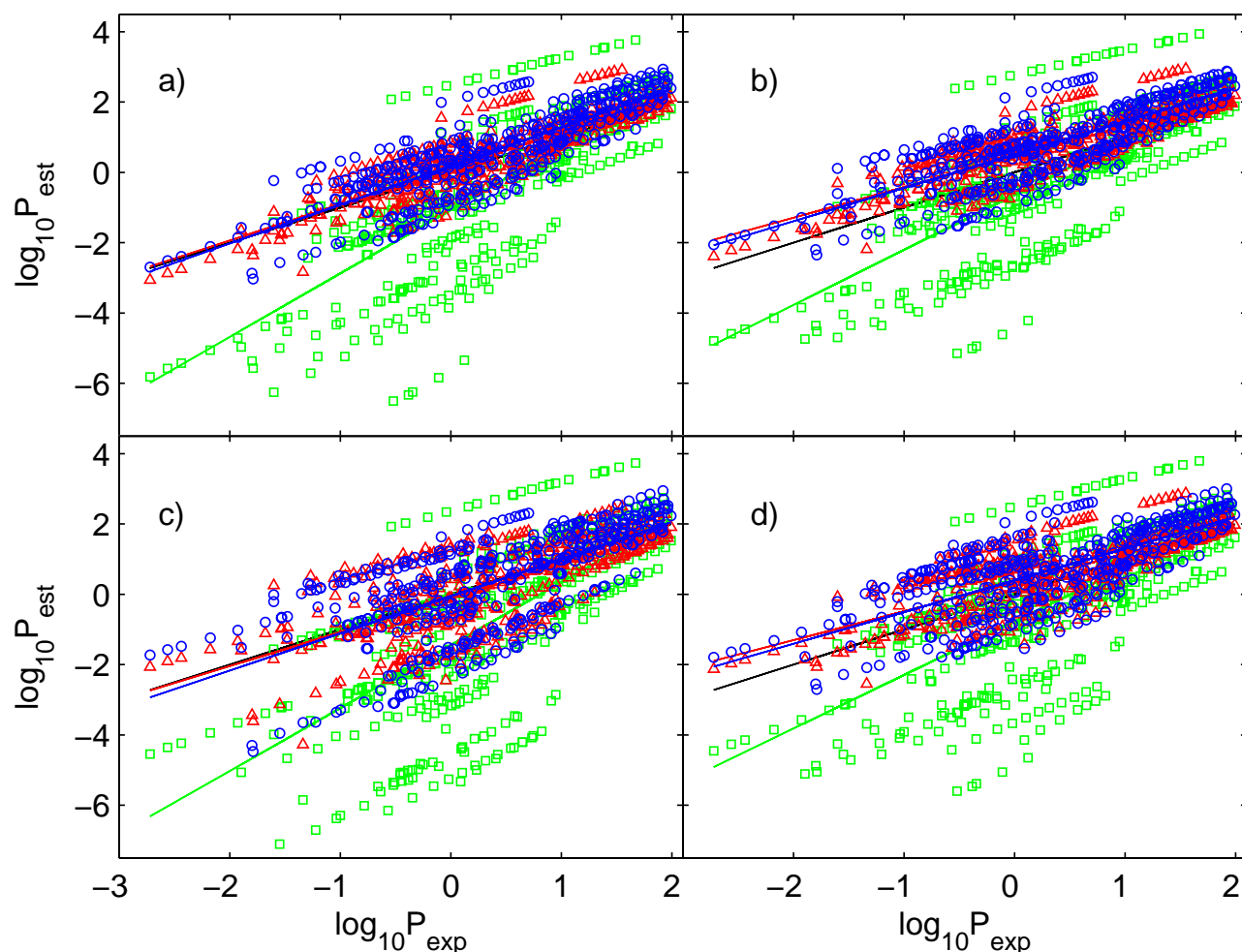


Fig. 1. P_{est} vs. P_{exp} plots for the 12 combined methods applied to Test Set 2. Each panel uses a different vapour pressure equation: (a) N-VP, (b) MY, (c) BV and (d) BK with the symbols colour-coded according to the boiling point estimation method used: Blue: N-Tb, Red: SB and Green: JR; where N-Tb= Nannoolal, SB= Stein and Brown, JR= Joback T_b estimation methods. The coloured lines are regression lines for each dataset (for coefficients see Table 6) and the black line is $X=Y$. The key to the vapour pressure equations is: N-VP= Nannoolal equation, BK= Baum equation with $\Delta S_{\text{vap}}=K_f R \ln(82.06 \cdot T_b)$, BV= Baum equation with Vetere equations, and MY= the Myrdal and Yalkowsky equation.

vapour pressures for the N-Tb/N-VP, N-Tb/M-VP and N-Tb/N-Sim methods shown in Fig. 4 are dominated by the errors in the T_b estimation method. The errors in predicting T_b dominate any differences in performance between these three vapour pressure equations leading to the greatly simplified N-Sim model giving similar results to the full N-VP method.

Myrdal and Yalkowsky (1997) reported that their method (MY) fitted their experimental database with an average factor error (as defined by Camredon and Aumont, 2006) of 1.62. Camredon and Aumont (2006) reported good results using the JR T_b estimation method with the MY vapour pressure equation, and did not report any significant bias in T_b from using the JR estimation method in contrast to the results noted above. This may be because most of the compounds in

their database had experimental T_b values below 550 K. They reported an average factor error of about 2 for all compounds and 3 for di- and tri-functionalised compounds. Nannoolal et al. (2008) claim a very low average factor error (<1.1) for a test set of 396 compounds using N-VP equation. Pankow and Asher (2008) quote an average error factor of about 2.2 for the SIMPOL.1 method. By contrast the average error factor for Test Set 2, calculated from the MBE values in Table 5, was 2.7 for T_b by N-Tb and vapour pressure estimated by the N-VP equation; 3.5 for T_b by N-Tb, with the MY vapour pressure equation; 5.2 for the SIMPOL.1 method; and 12.3 for T_b by JR coupled with the MY vapour pressure equation. These error factors are significantly higher than those quoted by the various authors for their respective methods and highlight how demanding the calculation of very low

Table 5. Results (MBE values^a) for the different vapour pressure estimation methods against Test Set 2.

No	Compound	Vapour pressure estimation method ^b																
		1	2	3	4	5	6	7	8	9	10	11	12	13 ^c	14 ^f	15	16	17
1	1,2-Pentanediol	0.197	0.492	-0.153	-0.007	0.318	-0.395	-0.084	0.274	-0.517	0.267	0.564	-0.085	-0.181	-0.335	0.072	-0.013	0.503
2	1,2,3-Trihydroxypropane	0.001	1.300	0.226	-0.008	1.338	0.225	-0.167	1.335	0.097	0.391	1.611	0.601	0.616	0.426	-0.217	-0.456	1.187
3	1,4-Butanediol	0.173	0.960	0.822	-0.011	0.858	0.706	-0.153	0.812	0.645	0.177	0.986	0.844	0.716	0.418	0.095	-1.117	0.622
4	2-(Methylamino)ethanol	0.117	0.807	0.728	-0.142	0.636	0.548	-0.301	0.572	0.475	-0.015	0.723	0.639	0.583	0.218	-0.180	-0.119	-0.119
5	2,4-Pentanediol	0.539	0.761	-0.307	0.396	0.640	-0.538	0.637	-0.659	0.642	0.864	-0.202	-0.337	0.056	0.440	0.442	0.786	0.442
6	2,6-Dinitrotoluene	0.173	0.245	-2.754	-0.034	0.042	-3.000	0.187	0.264	-2.956	0.164	0.236	-2.646	-0.212		-0.296	-0.280	-1.063
7	2-Aminoethanol	0.031	1.044	0.853	0.081	1.122	0.928	-0.131	1.052	0.835	-0.089	0.979	0.779	1.189		-0.458	-0.122	0.015
8	2-Aminonitrobenzene	-0.263	0.179	-1.185	0.087	0.509	-0.792	0.140	0.584	-0.802	0.144	0.556	-0.709	0.447		-0.104	-0.151	-0.543
9	2-Chloro-3,5-dimethoxy-4-HBA ^c	-1.345	-0.899	-2.464	-1.112	-0.678	-2.185	-2.671	-2.085	-4.144	-0.774	-0.369	-1.771			-1.545	-1.164	-1.909
10	2-Chloropropionic acid	0.003	0.206	0.216	0.406	0.594	0.604	-0.418	-0.160	-0.147	0.120	0.322	0.332			-0.005	-0.074	0.110
11	2-Hydroxybenzoic acid	0.201	-0.918	-2.024	0.723	-0.311	-1.322	-0.112	-1.483	-2.862	0.698	-0.314	-1.296	-1.656		0.133	-0.060	-0.015
12	2-Phenylbromide-TEGMMME ^c	0.912	0.728	-0.597	1.086	0.903	-0.403	1.396	1.219	-0.059	0.966	0.777	-0.571			0.610	0.540	0.319
13	3-(2-Methoxyphenyl)PA ^c	-0.103	-0.004	-1.002	0.730	0.816	-0.048	-0.698	-0.574	-1.835	0.522	0.611	-0.281	-0.482		-0.130	-0.696	-0.106
14	3-(3,4-Dimethoxyphenyl)PA ^c	0.136	0.415	-1.137	1.193	1.426	0.137	-0.292	0.045	-1.851	1.016	1.256	-0.068	-0.324		0.143	-0.469	0.156
15	3,4-Dichloronitrobenzene	0.007	0.056	-1.133	-0.111	-0.058	-1.335	0.023	0.076	-1.242	0.047	0.097	-1.112			-0.116	-0.062	-0.274
16	3,4-Dihydroxychlorobenzene	-0.421	-0.063	-0.795	-0.739	-0.346	-1.146	-1.446	-0.960	-1.957	-0.384	-0.026	-0.754			-0.505	-0.623	-0.967
17	3,5-di-tert-Butylcatechol	-0.780	-0.708	-3.403	-0.890	-0.815	-3.584	-1.679	-1.586	-5.127	-0.261	-0.196	-2.592	-2.540		-0.888	-0.883	-0.974
18	3,7-Dimethyl-7-hydroxyoctanal	0.477	0.366	-0.897	0.213	0.091	-1.295	-0.052	-0.193	-1.821	0.695	0.586	-0.641	-0.623	0.476	0.458	0.183	0.778
19	3-Chloro-2,6-dimethoxyphenol	-0.578	-0.283	-1.064	-0.686	-0.378	-1.191	-1.801	-1.393	-2.474	-0.459	-0.168	-0.934			-0.818	-0.404	-0.947
20	3-Chloroaminobenzene	0.285	0.310	0.232	0.258	0.285	0.202	0.305	0.333	0.246	0.265	0.292	0.210			0.230	0.292	0.148
21	3-Hydroxypropanenitrile	-0.183	0.551	0.354	-0.145	0.614	0.411	-0.433	0.450	0.217	0.008	0.721	0.529			0.197	-0.335	0.031
22	3-Nitro-3-(4-nitrophenyl)-pentane	0.000	0.182	-3.667	0.294	0.470	-3.146	0.558	0.733	-3.005	0.390	0.562	-2.946	-0.234		-0.068	-0.042	-0.583
23	3-Nitrobenzoic acid	0.064	0.229	-1.859	0.766	0.913	-0.928	-0.635	-0.422	-3.163	0.644	0.793	-1.060	-0.002		-0.031	-0.284	-0.208
24	3-Nitrophenol	1.414	0.857	-0.241	1.362	0.759	-0.442	0.952	0.184	-1.399	1.438	0.863	-0.272	1.852		1.359	1.378	1.080
25	4-Aminobenzoic acid	0.871	1.151	0.252	1.569	1.814	1.030	0.372	0.727	-0.421	1.318	1.575	0.753	0.990		0.845	0.673	0.798
26	Anisaldehyde	0.348	0.562	0.761	0.356	0.578	0.784	0.478	0.705	0.914	0.462	0.676	0.875	-0.137		0.331	0.278	0.062
27	Benzyl salicylate	-0.359	-0.959	-3.229	-0.137	-0.721	-2.876	-1.466	-2.253	-5.234	0.160	-0.386	-2.388	-1.856		-0.390	0.057	-0.635
28	Dibutyl phthalate	-0.273	0.072	-2.819	0.546	0.852	-1.656	0.596	0.914	-1.740	0.331	0.647	-1.936	-0.488		-0.213	-0.644	-0.693
29	Ethyl vanillin	0.263	0.327	-0.223	0.142	0.210	-0.374	-0.830	-0.740	-1.516	0.298	0.363	-0.192	-0.966		0.063	0.237	-0.357
30	Eugenol	-0.317	-0.239	-0.757	-0.513	-0.430	-0.983	-1.497	-1.388	-2.118	-0.377	-0.297	-0.826	-1.523		-0.480	-0.325	-0.765
31	Glutaric acid	0.799	0.920	0.108	1.574	1.678	0.982	0.399	0.549	-0.467	1.177	1.291	0.532	0.301	0.106	0.380	0.184	0.526
32	Glycerine carbonate	-0.044	-0.563	2.338	-0.175	-0.730	2.355	-0.704	-1.387	2.255	-0.378	-0.100	2.578			0.249	-0.486	0.542
33	Heliotropin	0.458	0.330	0.261	0.518	0.387	0.316	0.661	0.527	0.455	0.529	0.462	0.147			0.342	0.299	0.246
34	Isoamyl salicylate	-0.465	-1.186	-2.784	-0.341	-1.060	-2.626	-1.375	-2.324	-4.442	-0.197	-0.888	-2.383	-2.001		-0.504	-0.040	-0.835
35	Ketol	-0.637	-0.731	-2.891	-1.333	-1.442	-3.920	-1.842	-1.972	-5.005	-0.429	-0.519	-2.560	-0.985	-0.670	-1.050	-0.571	-0.321
36	Methyl anthranilate	-0.182	-0.101	-0.594	-0.199	-0.116	-0.626	-0.065	0.021	-0.507	-0.186	-0.104	-0.607	-0.315		-0.480	-0.201	-0.633
37	Musk ambrette	-1.181	-0.741	-5.805	-0.802	-0.384	-4.985	-0.472	-0.057	-4.814	-0.584	-0.186	-4.538	-1.825		-1.517	-1.374	-2.202
38	N-methyldiethanolamine	0.179	0.313	0.221	-0.128	0.023	-0.081	-0.255	-0.086	-0.203	0.228	0.364	0.270	-0.111		-0.100	-0.085	0.575
39	p-Acetylanisole	0.212	0.567	0.344	0.422	0.772	0.552	0.548	0.901	0.680	0.537	0.873	0.662	0.225		0.374	0.651	0.201
40	Phloroglucinol	1.527	1.943	1.116	1.566	1.996	1.138	1.117	1.644	0.584	1.661	2.070	1.256	-0.225		1.486	1.430	0.876
41	Pimelic acid	0.590	0.689	-0.536	1.570	1.652	0.639	0.292	0.411	-1.081	1.192	1.281	0.184	0.248	-0.458	0.250	0.249	0.451
42	Pinonaldehyde	0.572	0.240	-0.359	0.643	0.309	-0.290	0.903	0.569	-0.036	0.717	0.391	-0.192	-0.319	0.578	0.909	0.967	0.804
43	Tetraethylene glycol	0.038	-0.166	-1.172	-0.013	-0.228	-1.289	-0.218	-0.463	-1.687	0.454	0.261	-0.682	-0.659		-0.357	1.102	0.301
44	Triacetin	0.494	0.720	-1.058	1.163	1.365	-0.216	1.208	1.415	-0.231	1.086	1.292	-0.317	0.047		0.629	0.608	0.401
45	Triethylene glycol dinitrate	0.457	0.167	-4.075	1.319	1.071	-2.476	1.594	1.351	-2.239	1.157	0.897	-2.801	0.286		0.304	0.436	-0.021
Bias		0.10	0.23	-0.94	0.25	0.39	-0.72	-0.17	-0.03	-1.34	0.37	0.50	-0.57	-0.29	0.07	0.02	0.00	-0.06
St _D		0.56	0.66	1.60	0.74	0.80	1.53	0.93	1.05	1.83	0.57	0.63	1.37	0.94	0.46	0.62	0.59	0.74

^a MBE=(1/n)∑[log₁₀ p_{est}−log₁₀ p_{exp}].

^b Vapour pressure estimation methods: (1) N-Tb/N-VP; (2) SB/N-VP; (3) JR/N-VP; (4) N-Tb/BK; (5) SB/BK; (6) JR/BK; (7) N-Tb/BV; (8) SB/BV; (9) JR/BV; (10) N-Tb/MY; (11) SB/MY; (12) JR/MY; (13) SIMPOL.1; (14) CM method; (15) N-Tb/N-Sim; (16) N-Tb/M-VP; (17) N-Tb/LKA. For models 1–12, and 15–17 the first term is the key for the T_b estimation method (N-Tb= Nannoolal, SB= Stein and Brown, JR= Joback) and the letters after the slash are the key to the vapour pressure equation. N-VP= Nannoolal equation, BK= Baum equation with ΔS_{vap}=K_fRln(82.06·T_b), BV= Baum equation with Vetere equations, MY= the Myrdal and Yalkowsky equation, M-VP= the Moller equation, N-Sim= the simplified Nannoolal vapour pressure equation (see text), LKA= the Lee-Kesler equation with critical properties by the Ambrose method, and CM= method of Capouet and Muller.

^c Full names: compounds No. 9=2-chloro-3,5-dimethoxy-4-hydroxybenzaldehyde; No. 12=(2-phenylbromide)-triethyleneglycol-monomethylether; No. 13=3-(2-methoxy phenyl)propionic acid; No. 14=(3,4-dimethoxyphenyl)propionic acid.

^d St_D= standard deviation.

^e The SIMPOL.1 method was applicable to 36 compounds.

^f The CM method was applicable to 9 compounds.

vapour pressures for multifunctional compounds is for any estimation method. As discussed below the compounds in Test Set 2 are some 100–1000 times more volatile than the compounds expected to condense into organic aerosol (OA) so these errors are very likely to get worse when applied to atmospherically relevant partitioning.

4 The sensitivity of the partitioning of compounds into organic aerosol (OA) to vapour pressure values

4.1 An atmospherically relevant example

To represent the partitioning of semi-volatile organic compounds, it is convenient to use the volatility binning principle of Donahue et al. (2006) with the aim of modelling amounts of OA typical of moderately polluted ambient conditions (about 10 μg m⁻³ see their Fig. 1a). Donahue et al.

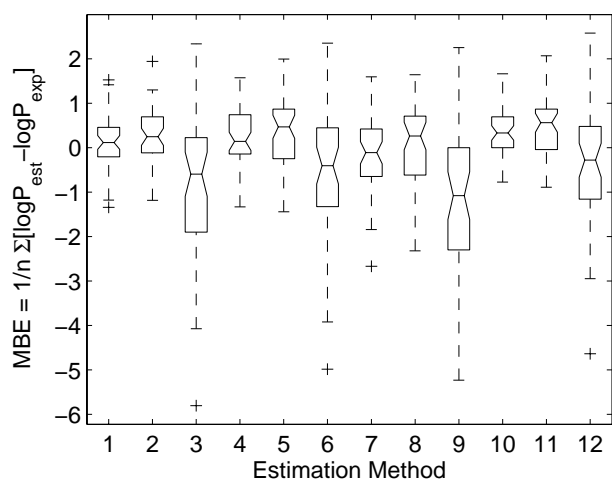


Fig. 2. Box-whisker plots of the MBE values for predicted vapour pressures of Test Set 2. The 12 combined estimation methods are: (1) N-Tb/N-VP; (2) SB/N-VP; (3) JR/N-VP; (4) N-Tb/BK; (5) SB/BK; (6) JR/BK; (7) N-Tb/BV; (8) SB/BV; (9) JR/BV; (10) N-Tb/MY; (11) SB/MY; (12) JR/MY. The first term is the key for the T_b estimation method (N-Tb= Nannoolal, SB= Stein and Brown, JR= Joback) and the letters after the slash are the key to the vapour pressure equation. N-VP= Nannoolal equation, BK= Baum equation with $\Delta S_{\text{vap}} = K_f R \ln(82.06 \cdot T_b)$, BV= Baum equation with Vetere equations, and MY= the Myrdal and Yalkowsky equation.

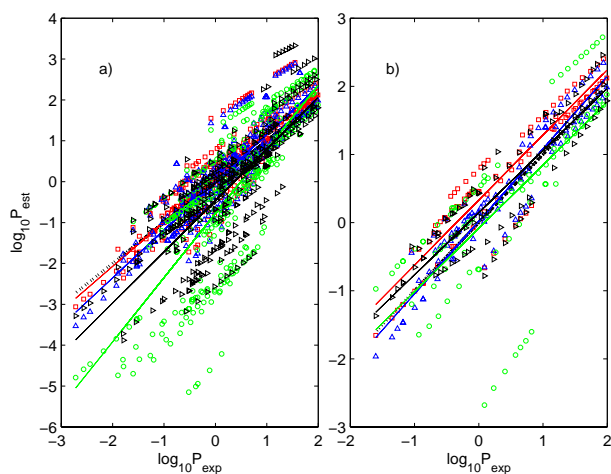


Fig. 3. P_{est} vs. P_{exp} plots comparing the predictions of SIMPOL.1 and the Capouet and Muller (CM) method against selected combined methods for relevant compounds. The same symbol colour is used for each combined method in both panels: N-Tb/N-VP (Red), N-Tb/M-VP (Blue) and JR/MY (Green). In panel (a) the SIMPOL.1 method (in Black) was compared to these methods for 36 compounds; and in panel (b) the CM method (Black) was compared for 9 compounds. The coloured lines are regression lines for each dataset (for coefficients see Table 6) and the $X=Y$ line is given by a thicker black dashed line. The key for the combined methods are: T_b estimation methods: N-Tb= Nannoolal, and JR= Joback; vapour pressure equations: N-VP= Nannoolal equation, M-VP= Moller equation and MY= the Myrdal and Yalkowsky equation.

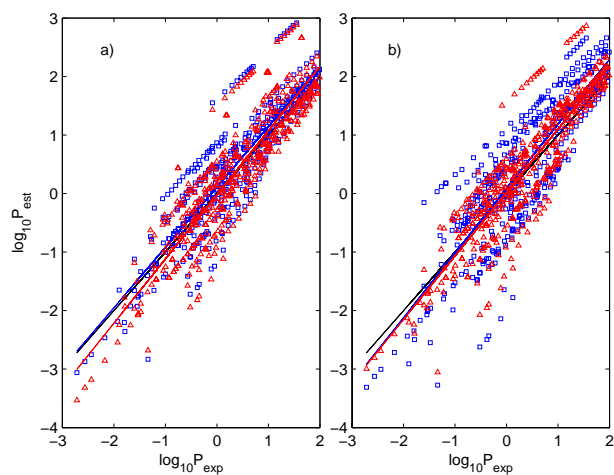


Fig. 4. P_{est} vs. P_{exp} plots comparing the predictions of some vapour pressure methods using the best T_b estimation method (the N-Tb or Nannoolal method- see text). In panel (a) the N-Tb/N-VP (Blue) combined method is compared to N-Tb/M-VP (Red) method. In panel (b) the N-Tb/LKA (Blue) combined method is compared to N-Tb/N-Sim (Red) method. The coloured lines are regression lines for each dataset (for coefficients see Table 6) and the black line is $X=Y$. The key for the vapour pressure equations is: N-VP= Nannoolal equation; M-VP= Moller equation; LKA= the Lee-Kesler equation with critical properties (calculated from N-Tb values) by the method of Ambrose; N-Sim= simplified Nannoolal equation (see text).

(2006) relate the formation of OA to the volatility of the condensing species using a saturated vapour density derived from the vapour pressure of each component, i :

$$C_i^* = \frac{10^6 M_i \gamma_i p_i^0}{RT}, \quad (27)$$

where γ_i is the activity coefficient (here assumed to be unity), M_i is the molecular weight and p_i^0 is the saturated liquid vapour pressure in atmospheres (sub-cooled if necessary) of component i ; R is the gas constant ($=8.2057 \times 10^{-5} \text{ atm m}^3 \text{ mol}^{-1} \text{ K}$) and the resulting C_i^* value is in $\mu\text{g m}^{-3}$. They propose binning the atmospheric compounds based upon the logarithm of their C_i^* value. In Fig. 1a of Donahue et al. (2006), where the formation of $10.6 \mu\text{g m}^{-3}$ of OA is modelled; components in bin 1 ($\log_{10} C_i^* = 1$ or $C_i^* = 10$) partition equally between the condensed and vapour phases while those in bin 0 largely condense. Hence for an example where $10.6 \mu\text{g m}^{-3}$ of OA is formed by partitioning then the maximum sensitivity of the amount of OA formed to vapour pressure values will be for those compounds in bins 0, +1 and +2. Compounds in higher bins only partition to a small extent and compounds in lower bins almost completely partition to the condensed phase so the amount of OA formed will be relatively insensitive to changes in vapour pressures for these compounds. However if changes

in vapour pressures result in compounds changing bins, then a much greater impact upon the amount of OA formed can occur. For example; if a compound in bin +3 (where it would have little impact upon the amount of OA formed) has its vapour pressure reduced by a factor of 100 this will cause it to switch to bin +1 where it will have a large effect upon the amount of OA formed. It is clear from Figs. 1 and 2 and Table 5 that changes in vapour pressures sufficient to cause compounds to change bins are quite common among the compounds of Test Set 2, with a potentially substantial impact upon the amount of OA formed.

The 45 multifunctional compounds can be assigned to bins using their C_i^* values as calculated using Eq. (27). All calculations were performed at a temperature of 25°C (298.15 K). Experimental vapour pressure data (subcooled liquid data where appropriate; see Table 3) was extrapolated/interpolated to 298.15 K by fitting to the two coefficient Antoine equation:

$$\ln(p_i^0) = A + B/T. \quad (28)$$

For the majority of compounds an extrapolation to 298.15 K was required. The two coefficient Antoine is known to give a reasonable approximation to known vapour pressures over small temperature intervals and is thus better at extrapolating vapour pressures than more complex forms with extra coefficients (Poling et al., 2001). These values at 298.15 K will be referred to as “experimentally derived” vapour pressures and will form the base case for the partitioning calculations. C_i^* values were calculated to give the following distribution of the 45 multifunctional compounds in the Donahue bins:

Bin +6, 3 compounds; +5, 9 compounds; +4, 12 compounds; +3, 9 compounds; +2, 7 compounds; +1, 4 compounds; 0, 1 compound.

Only one compound (phloroglucinol) is in bin 0, and 4 compounds (4-aminobenzoic acid, 3-(3,4-dimethoxyphenyl)propionic acid, pimelic and glutaric acids) are in bin +1. For the example where $10.6 \mu\text{g m}^{-3}$ of OA is formed, only phloroglucinol would show a high proportion of condensation while the other four compounds would show roughly equal partitioning between the condensed and vapour phases. All compounds, outside of these five, have been assigned to bins where only a small proportion (for many compounds, a very small proportion) of the compound would partition into the condensed phase. It should therefore be noted that the compounds of Test Set 2, despite their selection for multifunctionality and very low experimental vapour pressures, are still about 100–1000 times more volatile than the compounds believed to form OA under typical ambient conditions.

Table 6. The regression coefficients for the datasets.

VP Estimation Method ^a	Regression Coeffs. ^b	R^2
Fig. 1; $n=45$		
N-Tb/N-VP	1.0316X+0.1127	0.798
SB/N-VP	1.1018X+0.1800	0.7153
JR/N-VP	1.8064X−1.0682	0.6106
N-Tb/MY	0.8696X+0.4419	0.7332
SB/MY	0.9462X+0.5077	0.6614
JR/MY	1.5643X−0.6367	0.5845
N-Tb/BV	0.9669X−0.1209	0.5359
SB/BV	1.0566X−0.0548	0.4633
JR/BV	1.7903X−1.4472	0.4954
N-Tb/BK	0.8232X−0.3287	0.5922
SB/BK	0.9018X+0.3960	0.5351
JR/BK	1.5216X−0.7733	0.5147
Fig. 3; $n=36$		
N-Tb/N-VP	1.0916X+0.0973	0.8106
N-Tb/M-VP	1.1632X−0.0076	0.8084
JR/MY	1.5681X−0.7746	0.6033
SIMPOL.1	1.2567X−0.4538	0.6331
Fig. 3; $n=9$		
N-Tb/N-VP	0.9745X+0.3363	0.8661
N-Tb/M-VP	1.0380X+0.0450	0.9182
JR/MY	0.9283X−0.0851	0.5168
CM	0.9121X+0.1452	0.8402
Fig. 4; $n=45$		
N-Tb/N-VP	1.0278X+0.1174	0.7938
N-Tb/M-VP	1.1019X−0.0273	0.7955
N-Tb/LKA	1.1031X+0.0864	0.7410
N-Tb/N-Sim	1.080X+0.0214	0.7976

^a The first term is the key for the T_b estimation method (N-Tb= Nannoolal, SB= Stein and Brown, JR= Joback) and the letters after the slash are the key to the vapour pressure equation. N-VP= Nannoolal equation; BK= Baum equation with $\Delta S_{vap}=K_f R \ln(82.06 \cdot T_b)$; BV= Baum equation with Vetere equations; MY= the Myrdal and Yalkowsky equation; M-VP= Moller equation; LKA= the Lee-Kesler equation with critical properties (calculated from N-Tb values) by the method of Ambrose; and N-Sim= simplified Nannoolal equation.

^b Equation of line $\log_{10} P_{est}=A \log_{10} P_{exp}+B$, with $X=\log_{10} P_{exp}$ and A, B are the coefficients given in the table.

4.2 Partitioning of multifunctional compounds into OA under typical ambient conditions: the effect of changes in vapour pressure values

The partitioning model follows the approach described in Barley et al. (2009). This approach is based upon earlier models, e.g. Pankow (1994) with modification. The semi-volatile compounds are partitioned according to their saturation concentration (C_i^*) value:

$$C_i^* = \frac{10^6 \gamma_i P_i^0}{RT}. \quad (29)$$

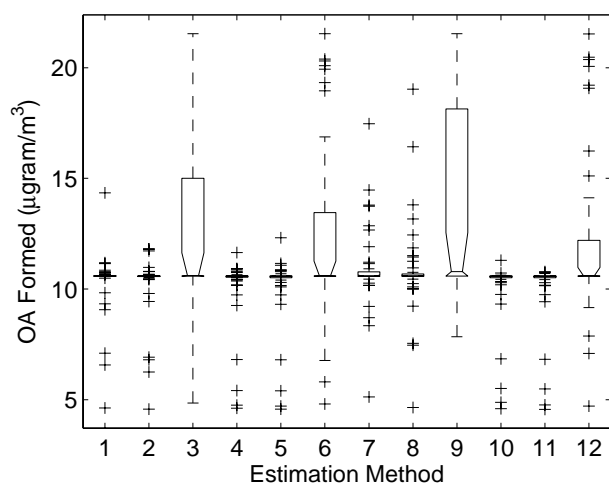


Fig. 5. Box-whisker plots of the predicted amount of OA formed by the partitioning model using estimated vapour pressures for a single compound from Test Set 2. The base case uses experimentally derived vapour pressures for all 45 compounds and gives $10.6 \mu\text{g m}^{-3}$ of OA. The vapour pressure of each compound in sequence is then changed to a value estimated by one of the following 12 methods and the amount of OA recalculated. The 12 combined estimation methods are: (1) N-Tb/N-Vp; (2) SB/N-Vp; (3) JR/N-Vp; (4) N-Tb/BK; (5) SB/BK; (6) JR/BK; (7) N-Tb/BV; (8) SB/BV; (9) JR/BV; (10) N-Tb/MY; (11) SB/MY; (12) JR/MY. The first term is the key for the T_b estimation method (N-Tb= Nannoolal, SB= Stein and Brown, JR= Joback) and the letters after the slash are the key to the vapour pressure equation. N-Vp= Nannoolal equation, BK= Baum equation with $\Delta S_{\text{vap}}=K_f R \ln(82.06 \cdot T_b)$, BV= Baum equation with Vetere equations, and MY= the Myrdal and Yalkowsky equation.

Where p_i^0 is the saturated vapour pressure of component i in atmospheres; R , T and γ_i have the same meaning as in Eq. (27); C_i^* is the saturation concentration in $\mu\text{mol m}^{-3}$. This can be converted to the C_i^* (in $\mu\text{g m}^{-3}$) described by Donahue (Eq. 27) by multiplying by M_i .

The amount of condensed material (C_{OA}) is then calculated by summing over all components i ensuring mole balance between the two phases for each component considered. Defining a partitioning coefficient ξ_i for compound i given its saturation concentration C_i^* (Eq. 29).

$$\xi_i = \left(1 + \frac{C_i^*}{C_{\text{OA}}}\right)^{-1}, \quad (30)$$

where both C_i^* and C_{OA} have units of $\mu\text{mol m}^{-3}$. The total number of moles of organic aerosol is the sum of the products of the individual component concentrations (C_i) and their partitioning coefficient (ξ_i):

$$C_{\text{OA}} = \sum_i C_i \xi_i. \quad (31)$$

This calculation provides the amount of each component in the condensed phase in $\mu\text{mol m}^{-3}$ and is readily converted

into mass based amounts by multiplying by the appropriate molecular weight. Summing the mass based condensed quantities for all the compounds provides the total condensed OA in mass based units.

In this example the C_i^* for all 45 multifunctional compounds were calculated using Eq. (29) and experimentally derived vapour pressures at 298.15 K. All compounds were assigned the same individual component concentration (C_i) and this was set to a value ($0.02666 \mu\text{mol m}^{-3}$) such that the partitioning model converged ($C_{\text{OA}}=0.062715 \mu\text{mol m}^{-3}$) to give $10.6 \mu\text{g m}^{-3}$ of OA. This formed the base case for the effect of changing vapour pressures on the partitioning of compounds to OA.

If the partitioning is recalculated with the experimentally derived vapour pressures replaced by estimated values then the compounds may move between the Donahue bins (see end of Sect. 4.1) and the amount of predicted OA varies dramatically (between 0 and $110 \mu\text{g m}^{-3}$) depending upon which estimation method is used. If the experimentally derived vapour pressures are all doubled the amount of predicted OA falls to zero (from $10.6 \mu\text{g m}^{-3}$). This demonstrates the sensitivity of OA formation to errors in estimated vapour pressures.

A large part of this extreme sensitivity of the amount of OA formed to the vapour pressure values used in the partitioning calculation is a reflection of the fact that this set of multifunctional compounds is too volatile to consistently form atmospherically relevant amounts of OA. A different approach is required to compare the effect of the 12 vapour pressure estimation methods shown in Figs. 1 and 2 on the amount of OA formed in this example. The base case described above was repeatedly run with a single vapour pressure value replaced by an estimated value. This was performed for each compound in sequence to give a distribution of 45 OA masses that can then be represented on a box-whisker plot. The results for all 12 estimation methods are shown in Fig. 5. The striking feature about this plot is that despite the fact that only 1/45th of the total number of moles in the mixture is assigned the estimated vapour pressure, some compounds cause the predicted amount of OA to change by a factor of two or more; and this is true for all the estimation methods, even those that give the most accurate predictions. In Fig. 5 the vast majority of compounds had little effect on the amount of OA formed because they are too volatile. This gives a very small box and a very similar distribution of outliers for all the estimation methods except those using T_b by the JR method. As mentioned above this method significantly overestimates T_b leading to many compounds that would be assigned (on the basis of their experimentally derived vapour pressures) to bins too high to significantly impact upon the amount of OA formed moving two or more bins down into the range where they do significantly affect the amount of OA predicted.

To better differentiate between the methods that did not use T_b by the JR method, the results for the 12 least volatile

compounds, based upon their experimentally derived vapour pressures (bins 0 to +2: see above) were replotted in Fig. 6. This plot does allow some differentiation between the predictions of the non-Joback methods but also emphasizes how important it is to obtain accurate vapour pressures for those compounds that are in the critical bins (in the example used here: bins 0 to +2) where ξ_i (see Eq. 30) is in the range 0.05–0.95.

5 Conclusions

The estimation method reported in Nannoolal et al. (2004) provide the most accurate T_b values. Stein and Brown (1994) provided the second best estimation method for T_b . The prediction of vapour pressures for the 45 multifunctional compounds of Test set 2 showed that the method of Nannoolal et al. (2008) and the method of Moller et al. (2008) were better than the other vapour pressure methods studied when used with the Nannoolal et al. (2004) T_b estimation method. However the results for both these vapour pressure equations (and the simplified form of the Nannoolal vapour pressure equation) are so similar that it is not possible, on the basis of this work, to select between them. The errors in the estimation of T_b dominate the smaller differences between the vapour pressure equations. The Moller et al. (2008) vapour pressure equation may be preferred to the other methods because the extra term for alcohols and acids potentially allows improved accuracy for these atmospherically important compounds (Bilde et al., 2003). This work highlights that, for a method that combines T_b estimation with a vapour pressure equation, improvement in the T_b estimation method will yield the greatest improvement in model skill for atmospherically important compounds. However the database of experimental T_b values is limited and contains few multifunctional compounds of the type seen in Test set 2 because these compounds have sufficiently high T_b values that they decompose before the normal boiling point temperature is reached. To improve T_b estimation methods more boiling point data for relatively involatile multifunctional compounds are required. This could be provided by extrapolating experimental vapour pressure data up to atmospheric pressure using a reliable vapour pressure equation. These “pseudo-experimental boiling points” could be combined with true experimental boiling point values for less structurally complex compounds and used to improve the estimation of normal boiling points.

The sensitivity of the amount of OA formed to the vapour pressure values used in the partitioning calculation was investigated for an atmospherically relevant example from the literature. It was found that despite the selection of the multifunctional compounds for their low vapour pressures they were still some 100–1000 times more volatile than the compounds expected to contribute significantly to OA. The amount of OA predicted to form (and by extension its composition) was extremely sensitive to changes in vapour pressure

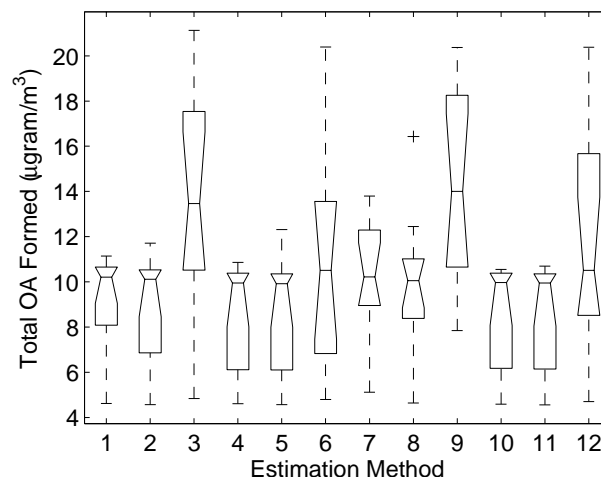


Fig. 6. Box-whisker plots of the predicted amount of OA formed by the partitioning model using estimated vapour pressures for a single compound from Test Set 2. This plot uses the same calculations to those used to generate Fig. 5 but only shows data for the 12 least volatile compounds (from bins 0, +1 and +2). The base case uses experimentally derived vapour pressures for all 45 compounds and gives $10.6 \mu\text{g m}^{-3}$ of OA. The 12 methods are: (1) N/N; (2) S/N; (3) J/N; (4) N/BK; (5) S/BK; (6) J/BK; (7) N/BV; (8) S/BV; (9) J/BV; (10) N/MY; (11) S/MY; (12) J/MY. The first letter is the key for the T_b estimation method (N= Nannoolal, S= Stein and Brown, J= Joback) and the letters after the slash are the key to the vapour pressure equation: N= Nannoolal equation, BK= Baum equation with $\Delta S_{\text{vap}} = K_f R \ln(82.06 \cdot T_b)$, BV= Baum equation with Vetere equations, MY= the Myrdal and Yalkowsky equation.

values (especially for the least volatile components) typical of the differences seen between estimation methods. The potential impact of errors in estimated vapour pressure values upon the amount and composition of predicted OA should not be underestimated.

Acknowledgements. This work was carried out within the UK NERC-funded “QUantifying the Earth SysTem” (QUEST) project under the “QUest Aerosol and Atmospheric Chemistry” (QUAAC) grant number NE/C001613/1) and EU-funded “European Integrated project on Aerosol Cloud Climate and Air Quality interactions” (EUCAARI) under contract number 036833-2.

Edited by: W. E. Asher

References

- 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.
- 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, 2005, <http://www.atmos-chem-phys.net/5/2497/2005/>.
- Barley, M., Topping, D. O., Jenkin, M. E., and McFiggans, G.: Sensitivities of the absorptive partitioning model of secondary organic aerosol formation to the inclusion of water, *Atmos. Chem. Phys.*, 9, 2919–2932, 2009, <http://www.atmos-chem-phys.net/9/2919/2009/>.
- Barsanti, K. C. and Pankow, J. F.: Thermodynamics of the formation of atmospheric organic particulate matter by accretion reactions - Part 1: aldehydes and ketones, *Atmos. Environ.*, 38, 4371–4382, 2004.
- Baum, E. J.: *Chemical property estimation-Theory and application*, Section 6.3., CRC Press, Boca Raton, FL, USA, 1998.
- Bencze, L., Raj, D. D., Kath, D., Oates, W. A., Singheiser, L., and Hilpert, K.: Thermodynamic properties and diffusion thermodynamic factors in B2-NiAl, *Metall. Mater. Trans. B.*, 35, 867–876, 2004.
- 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.
- Bloss, C., Wagner, V., Jenkin, M. E., Volkamer, R., Bloss, W. J., Lee, J. D., Heard, D. E., Wirtz, K., Martin-Reviejo, M., Rea, G., Wenger, J. C., and Pilling, M. J.: Development of a detailed chemical mechanism (MCMv3.1) for the atmospheric oxidation of aromatic hydrocarbons, *Atmos. Chem. Phys.*, 5, 641–664, 2005, <http://www.atmos-chem-phys.net/5/641/2005/>.
- Booth, A. M., Markus, T., McFiggans, G., Percival, C. J., McGillen, M. R., and Topping, D. O.: Design and construction of a simple Knudsen effusion mass spectrometer (KEMS) system for vapour pressure measurements of low volatility organics, *Atmos. Meas. Tech. Discuss.*, 2, 893–914, 2009, <http://www.atmos-meas-tech-discuss.net/2/893/2009/>.
- Cambridgesoft: *ChemDraw Ultra version 10.0*, Cambridge, MA, USA, 2005.
- Camredon, M. and Aumont, B.: Assessment of vapor pressure estimation methods for secondary organic aerosol modeling, *Atmos. Environ.*, 40, 2105–2116, 2006.
- 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, 2006, <http://www.atmos-chem-phys.net/6/1455/2006/>.
- 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.
- Clegg, S. L., Kleeman, M. J., Griffin, R. J., and Seinfeld, J. H.: Effects of uncertainties in the thermodynamic properties of aerosol components in an air quality model - Part 2: Predictions of the vapour pressures of organic compounds, *Atmos. Chem. Phys.*, 8, 1087–1103, 2008, <http://www.atmos-chem-phys.net/8/1087/2008/>.
- Copland, E. H. and Jacobson, N. S.: Thermodynamic activity measurements with Knudsen cell mass spectrometry (KCMS or KEMS), *Electrochem. Soc.-Interface*, 6, 28–31, 2001.
- Dekruif, C. G. and Vanginkel, C. H. D.: Torsion-weighting effusion vapor-pressure measurements on organic-compounds, *J. Chem. Thermodyn.*, 9, 725–730, 1977.
- Donahue, N. M., Robinson, A. L., Stanier, C. O., and Pandis, S. N.: Coupled partitioning, dilution, and chemical aging of semivolatile organics, *Environ. Sci. Technol.*, 40, 2635–2643, 2006.
- EPA: *Estimation Programs Interface Suite for Microsoft Windows v3.2*, United States Environmental Protection Agency, Washington, DC, USA, 2009.
- Faber, N. M. and Rajko, R.: How to avoid over-fitting in multivariate calibration - The conventional validation approach and an alternative, *Anal. Chim. Acta*, 595, 98–106, 2007.
- Fishtine, S. H.: Reliable latent heats of vaporization, *Ind. Eng. Chem.*, 55, 47–56, 1963.
- Goodman, B. T., Wilding, W. V., Oscarson, J. L., and Rowley, R. L.: Use of the DIPPR database for development of quantitative structure-property relationship correlations: Heat capacity of solid organic compounds, *J. Chem. Eng. Data*, 49, 24–31, 2004.
- Grain, C. F.: Vapor pressure, in: *Handbook of chemical property estimation methods*, edited by: Lyman, W. J., Roehl, W. F., and Rosenblatt, D. H., McGraw Hill, Chapter 14, 1982.
- Hallquist, M., Wangberg, I., and Ljungstrom, E.: Atmospheric fate of carbonyl oxidation products originating from α -pinene and Δ (3)-carene: Determination of rate of reaction with OH and NO₃ radicals, UV absorption cross sections, and vapor pressures, *Environ. Sci. Technol.*, 31, 3166–3172, 1997.
- Joback, K. G. and Reid, R. C.: Estimation of pure-component properties from group-contributions, *Chem. Eng. Commun.*, 57, 233–243, 1987.
- Johnson, D., Utembe, S. R., and Jenkin, M. E.: Simulating the detailed chemical composition of secondary organic aerosol formed on a regional scale during the TORCH 2003 campaign in the southern UK, *Atmos. Chem. Phys.*, 6, 419–431, 2006, <http://www.atmos-chem-phys.net/6/419/2006/>.
- Kapteina, S., Slowik, K., Verevkin, S. P., and Heintz, A.: Vapor pressures and vaporization enthalpies of a series of ethanolamines, *J. Chem. Eng. Data*, 50, 398–402, 2005.
- Koponen, I. K., Riipinen, I., Hienola, A., Kulmala, M., and Bilde, M.: Thermodynamic properties of malonic, succinic, and glutaric acids: Evaporation rates and saturation vapor pressures, *Environ. Sci. Technol.*, 41, 3926–3933, 2007.
- Lei, Y. D., Wania, F., Shiu, W. Y., and Boocock, D. G. B.: Temperature dependent vapor pressures of chlorinated catechols, syringols, and syringaldehydes, *J. Chem. Eng. Data*, 44, 200–202, 1999.
- Lyman, W. J.: *Environmental exposure from chemicals*, CRC Press, Boca Raton, FL, USA, vol 1, chapter 2, 1985.
- 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 Equilib.*, 269, 117–133, 2008.
- Pankow, J. F.: An absorption-model of gas-particle partitioning of organic-compounds in the atmosphere, *Atmos. Environ.*, 28, 185–188, 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, 2008, <http://www.atmos-chem-phys.net/8/2773/2008/>.
- Poling, B. E., Prausnitz, J. M., and O'Connell, J. P.: *Properties of gases and liquids*, Chapter 7, McGraw-Hill, New York, NY, USA, 2001.
- Prausnitz, J. M., Lichtenhaler, R. N., and de Azevedo, E. G.: *Molecular Thermodynamics of Fluid-Phase Equilibria*, Prentice-Hall, Englewood Cliffs, NJ, USA, 2nd edition, 1986.
- Reid, R. C., Prausnitz, J. M., and Poling, B. E.: *Properties of gases and liquids*, McGraw-Hill, New York, NY, USA, 1987.
- Ruzicka, V. and Domalski, E. S.: Estimation of the Heat-Capacities of Organic Liquids as a Function of Temperature Using Group Additivity, 2. Compounds of Carbon, Hydrogen, Halogens, Nitrogen, Oxygen, and Sulfur, *J. Phys. Chem. Ref. Data*, 22, 619–657, 1993.
- Saxena, P. and Hildemann, L. M.: Water-soluble organics in atmospheric particles: A critical review of the literature and application of thermodynamics to identify candidate compounds, *J. Atmos. Chem.*, 24, 57–109, 1996.
- Stein, S. E. and Brown, R. L.: Estimation of normal boiling points from group contributions, *J. Chem. Inf. Comp. Sci.*, 34, 581–587, 1994.
- Tao, Y. and McMurray, P. H.: Vapor-pressures and surface free-energies of C14-C18 monocarboxylic acids and C5-dicarboxylic and C6-dicarboxylic acids, *Environ. Sci. Technol.*, 23, 1519–1523, 1989.
- Verevkin, S. P., Wandschneider, D., and Heintz, A.: Determination of vaporization enthalpies of selected linear and branched C-7, C-8, C-9, C-11, and C-12 monoolefin hydrocarbons from transpiration and correlation gas-chromatography methods, *J. Chem. Eng. Data*, 45, 618–625, 2000.
- Vetere, A.: Methods to predict the vaporization enthalpies at the normal boiling temperature of pure compounds revisited, *Fluid Phase Equilib.*, 106, 1–10, 1995.
- Wexler, A. S. and Clegg, S. L.: Atmospheric aerosol models for systems including the ions H^+ , NH_4^+ , Na^+ , SO_4^{2-} , NO_3^- , Cl^- , Br^- , and H_2O , *J. Geophys. Res.-Atmos.*, 107, 4207, doi:10.1029/2001JD000451, 2002.