**Atmospheric
Chemistry
and Physics
Discussions**

Interactive comment on "EVAPORATION: a new vapor pressure estimation method for organic molecules including non-additivity and intramolecular interactions" *by* S. Compernolle et al.

S. Compernolle et al.

steven.compernolle@aeronomie.be

We thank all referees sincerely for their detailed comments and useful suggestions, which helped to improve the manuscript. Our responses are given below.

Answer for referee #1

General Points

1. *Two modelling methods for the calculation of (subcooled) liquid vapour pressures are introduced. First, a group-additive method with some additional second order descriptors applicable to non-functionalized hydrocarbons and to monofunctional compounds. Second, a method termed "full method" that is improved for the description of multifunctional compounds. It is not clear to the reader, which method is used as the "EVAPORATION" method. Do the authors apply both methods under this name depending on the compound type? Which method has been used for the comparison of monofunctional compounds with experimental data and other vapour pressure estimation methods in Section 5 and corresponding figures (e.g., Fig. 6)? I might have missed it, but it is also not clear to me, whether the authors used the full database including monofunctional compounds for the parameter fitting of the full method. A few sentences concerning these points would help the reader understand what has been done.*

The method termed "full method" is fitted using the full database, including the monofunctional compounds and the hydrocarbons. In the comparison with other methods (Fig. 5) and with experiment (Fig. 6) always this method has been used. Of course if people only need vapour pressures of monofunctional compounds, they could use the method optimized solely for hydrocarbons and monofunctional compounds.

In sec 4.3.3 the sentence is added: "These parameters are fitted to the data of all compounds (hydrocarbons, monofunctional and polyfunctional molecules)."

In sec 5.1, line 16, and in the captions of Figs. 5 and 6, "EVAPORATION" has been replaced by "EVAPORATION (the full model described in sec 4.3.3)".

2. *Related to the point above. If hydrocarbons and monofunctional compounds were*

also used to fit the parameters of the full method, how do the predictions of the two methods compare for monofunctional groups?

One can compare the different classes of monofunctional compounds of Tables 3 and 7 to this end. Difference in performance is slight for both methods. Although the overall performance of the full method is slightly less for hydrocarbons+monofunctional compounds, the difference is not large. We add in the text p 13252, line 18:

"For hydrocarbons and monofunctional compounds, the MAD and pred. MAD for hydrocarbons and monofunctional compounds is 0.085 and 0.087 respectively. This is only slightly higher than as obtained with the method of Section 4.2.2 (see Table 3)."

3. While the mathematical expressions for the calculation of vapour pressures with the proposed method are relatively simple, the correct assignment of the different descriptors to a given compound is likely the more complicated step. It might therefore be helpful if the authors present a few calculation examples illustrating the assignments of the descriptors to different compounds, perhaps in an appendix.

This is a good suggestion. We add in the supplementary material a list of examples, for hydrocarbons, monofunctional and polyfunctional molecules.

4. According to the ACP guidelines, SI units should be used wherever possible. The authors use three different units for pressure in the text and figures (atm, Torr, kPa). I consider this a minor issue, but worth mentioning. While Torr is an often encountered unit for vapour pressure data, there is no reason not to use units of Pa (or kPa).

We would like to keep the unit of atm, even if it is no SI unit, in the derivation in Section 4.1 (important also for the zero-point value in Tables 2 and 4) as it results in a particular simple form of the term A (Eq. (18)). However, we agree that the use of different units can be confusing. Therefore, we convert everywhere Torr (the main unit used in the discussion paper) to atm, and at the single occurrence of 1 kPa (following a citation of another work) we add "$(9.87 \cdot 10^{-3} \text{atm})$".

5. Given the applicability and predictability of the EVAPORATION method for a large variety of organic compounds, this method has the potential to become widely used in the community. Do the authors intend to make EVAPORATION available to the wider community, for example via a web-interface?

We are in discussion with profs. Wexler and Clegg to place EVAPORATION on the website E-AIM (http://www.aim.env.uea.ac.uk/aim/aim.php) where several models regarding aerosol thermodynamics are already present.

Specific Points

• *Abstract: "a method to predict vapour pressure p0". At least at this point it would be more correct to write "a method to predict pure compound (subcooled) vapour pressure p0", as to clearly distinguish it from the vapour pressure of compounds in mixtures or the vapour pressure of a solid. The same distinction should be made at the first occurrence of "vapour pressure" in the Introduction.*

We put "(subcooled) liquid pure compound vapour pressure " in the abstract and the introduction.

• *p. 13231, l. 16-19: The sentence at these lines could be understood as if there is no positive aspect of using boiling points with vapour pressure estimation methods. It should also be mentioned why some methods benefit from boiling point data as input.*

We changed the sentence into "This is an advantage if this boiling point is experimentally known, but it can contribute to the overall error if it has to be estimated."

• *p. 13232, l. 18: a pressure-temperature (p0(T)) correlation,", are these correlations e.g. in the form of Antoine Equation parameters? If so, it might be worth mentioning.*

We introduced "$(p^0(T))$ correlation -e.g.~an Antoine equation or a Wagner equation-"

• *p. 13233, l. 12: "Following groups of compounds can be distinguished:" Here either*

the rest of the sentence is missing or the whole sentence should be written differently to introduce the different compound classes discussed in the following Sections.

We added: "non-functionalized hydrocarbons, monofunctional compounds and poly-functional compounds."

• p. 13235, Eq. (1): No citation for the source of this equation is given in the text. It should also be mentioned that in Eq. (1) the melting point temperature (Tfus) is used instead of the (theoretically correct) temperature at the triple point.

We introduced a reference to the handbook of Prausnitz (1999), and the remark on triple point: "Note that in Eq. (1) T_{fus} is used instead of the (theoretically correct) triple point temperature, but this incurs little error."

• p. 13240, l. 1: "sublimation data of Soonsin et al. (2010) was selected, partly because their data on saturated solutions should be more reliable than data on solid particles themselves," This formulation is confusing and needs editing. If the saturated solutions data is more reliable than the sublimation data of solid particles, why would one use the sublimation data (of solid particles) and not the data from the saturated solutions. Of course, the sublimation vapour pressure of the solid particles and the vapour pressures of the saturated solutions should be equal, but that is not what the sentence is saying. What data of Soonsin et al. (2010) have been used, sublimation data of solids or saturated solution data?

Soonsin et al. (2010) gives two reasons why data on saturated solutions should be more reliable. First, the possible non-sphericity of solid particles gives uncertainty on the size and evaporation rate of the particles. Second and probably more importantly, evaporation rates (related to the vapour pressure) of solid particles are not constant in time. At the start they resemble the vapour pressure of the pure liquid, at the end that of the saturated solution. Their conclusion is that these solid particles probably contain liquid inclusions that have to evaporate first before solid state vapour pressure can be measured.

Thus, we chose for malonic, succinic and glutaric acid the saturated solution data. For oxalic acid, the saturated solution data referred to the dihydrate, hence here we chose the anhydrous solid state vapour pressure. We changed the sentence into:

"Second, the saturated solution data of Soonsin et al. (2010) was selected for malonic, succinic and glutaric acid, while we chose the solid particle data for oxalic acid. Their data on saturated solutions is more reliable than their data on solid particles themselves, as the possible non-sphericity of solid particles bears uncertainty on measured vapour pressure, and the non-constant evaporation rates, probably due to liquid inclusions, complicates their measurements. For oxalic acid the solid particle data was chosen, as the saturated solution data was for the dihydrate rather than anhydrous oxalic acid. Another reason to choose the data of Soonsin et al. (2010) was its consistency with the corresponding liquid vapour pressure data and the fusion enthalpy."

• p. 13241, l. 4: With Section 3, there is a somewhat sudden change in topic. It is not clear from the first sentence in this section, why there should be statistical evaluators discussed / introduced at this point. A sentence or two, introducing the idea and use of statistical evaluators later in the model description, would likely enhance the readability of and justification for this section.

At the start of Section 3, we write: "Before describing the method framework of EVAPORATION and the procedure to fit its parameters, we will describe here the statistical evaluators that will be used to report the performance of the EVAPORATION. They include the model bias or mean deviation (MD), the mean absolute deviation (MAD), indicating the ability of the model to fit the data, and the predicted MD and MAD, indicating the predictivity of the model. MD and MAD will also be used to report the performance of other vapour pressure estimation methods on the molecules in our database."

• p. 13243, l. 1: "Note that setting K = 1 returns the basic Clausius-Clapeyron equation under assumption of a temperature-independent enthalpy of vaporisation, valid only in

a small temperature interval." It might be worth mentioning that with K = 1, Eq. (13) is also known as the August Equation.

We introduced: "also known as the August equation".

• p. 13243, l. 7: It should be noted that "term A is directly related to the entropy of boiling" at 1 atm total pressure.

"at 1 atm total pressure" inserted.

• p. 13244, l. 5: At this point in the text, it is unclear for a reader what the authors mean by "first-order group contribution" and "second-order corrections". Please explain.

The sentence is added: "The first-order groups describe the molecule as a set of fragments (carbon atoms and functional groups), while the second-order groups take the environment of functional groups into account."

• p. 13244, l. 6: "and ck are descriptor values of the molecule. A descriptor is a property of the molecule that is readily obtained or calculated, e.g. the number of times a certain functional group occurs." These two statements can lead to some confusion concerning the meaning and use of the terms "descriptor" and "descriptor values" as later in the manuscript it seems that the descriptors are actually the accountable molecule properties listed in Tables 2 and 4, indexed by k (attributed with specific ak and bk values). See also the statement on line 2 of page 13245. Could the set of ck be described as a molecule-specific list (vector) with entries that relate to the number of times a certain "descriptor k" has to be accounted for, i.e., by evaluating column 1 of Tables 2 and 4?

We avoid the use of "number of times a certain descriptor has to be accounted for" in the general definition, as also non-integer and/or non-positive values for a descriptor are in principle possible. This is the case for $c_3 = t$, which can take on negative values.

Sentence has changed into:

"c_k are the values of a set of molecular descriptors. These descriptors are accountable

C7469

molecular properties, calculated from molecular structure information. An important example is the number of times a certain functional group occurs. "

• p. 13245, l. 22 (Eq. 24): While I agree with the authors' conclusion that the "#observables" in the definition of degree of freedom, df, should be related to the "(number of species where descriptor occurs)", I wonder why in a similar sense "#parameters" is not also replaced by #(parameters influencing species where descriptor occurs) in Eq. (24) instead of ignoring it.

Note first that on line 14, we changed the Raventos *et al.* (2010) definition from "df = #species − #parameters − 1" to "df = #species − #parameters". The confusion did arise because we overlooked that the "zero term" is not seen as a descriptor by Raventos *et al.* (2010).

In the discussion on significance, through Eq. (23), we considered only the total group contributions at 298 K, g_k and not a_k, b_k separately. Hence let us assume for simplicity in the following discussion that our method is a one temperature method (at 298 K) with parameters g_k, and each species corresponds to one observable.

Each parameter g_k corresponds to a descriptor (if one counts, as we do, the zero term also as a descriptor). In our definition of df, there is a separate equation for df per descriptor (or parameter). Hence there is no need to take the other parameters into account in Eq. (24). One could also consider this as the appropriate case when all other parameters $g_{l \neq k}$ are fixed, i.e. if they were as experimental values, and that only g_k is the parameter to be obtained. Then the problem reduces to the determination of a single parameter from a set of measurements, and df = #observables − 1 (the "1" because there is one parameter to determine).

To make clearer that in our definition df refers to a single descriptor, we changed the sentence on line 23 into

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

C7470

• *p. 13248, l. 17: Related to my general comments, in case the "Full method" is not recommended / used for the calculation of vapour pressures of monofunctional compounds, then this Section would probably be better named "Method for multifunctional compounds".*

The method can be and is used both for monofunctional and multifunctional compounds, hence we keep the title.

• *p. 13251, l. 15: "A molecule with two vicinal carbonyl groups will have a higher vapour pressure, because", higher vapour pressure than what?*

We changed the sentence into "A molecule with two carbonyl groups will have a higher vapour pressure if these groups are vicinal"

• *p. 13253, l. 14: replace "generally" by "typically" in the Section title.*

Done.

• *p. 13254, l. 14: "had to be below the MG", abbreviation MG is not defined.*

Sentence is changed into:

"Another requirement was that the temperature had to be below the critical temperature as estimated by the method of Marrero and Gani (2001) (MG)"

• *p. 13255, l. 12: "the predictions of EVAPORATION", maybe better: "the predicted MAD of EVAPORATION"*

We inserted "the predicted MAD of EVAPORATION is only slightly above the MAD from the fitting"

• *p. 13255, l. 19: why "ad-hoc modification"? delete "ad-hoc". Same thing on the following page.*

With the term "ad-hoc" we wanted to indicate that the modification is difficult to explain from a physico-chemical point of view. The term has been deleted.

• *p. 13265, Table 2: Related to the point about the term "descriptor", would it be correct to add "Descriptor, ck" as header of column 1? At the entry "Zero point" one could add "Zero point (c1 = 1).*

This has been done.

It should also be noted that the symbol "#" in column 1 stands for "number of".

This is done.

Functional groups: instead of "-OH (alcohol)" writing "-OH (hydroxyl)" would be better, as I believe this descriptor is also used for OH-groups in non-alcohols (e.g. citric acid), or not?

We have now, here and at other places in the text, replaced 'alcohol' by 'hydroxyl' if it referred to the functional group rather than the molecule.

• *p. 13267, Table 4: Column "Type": mentioning in a footnote again what the abbreviations lin, CL, and HB stand for would be helpful to read the table.*

This has been done.

• *p. 13270, Table 7: Naming of molecule classes. Could "Carbonyls" be replaced by "Aldehydes + Ketones", as carbonyl is a rather general group name that includes carboxylic acids.*

A correct and concise naming is difficult as we consider aldehydes and ketones always together, (given their very similar vapour pressures, and similar structure), but no name seems to exist to designate these two, and only these two. With carbonyls, we meant aldehydes and ketones only. We have noted that "Carbonyls" refer to aldehydes and ketones in Tables 2 and 7. Furthermore, with the "carbonyl group" we meant specifically the aldehydic or the ketone group, not the ester group or the carboxylic acid group. This is now mentioned clearly in section 4.2.2, p 13246, line 26: "Note that with the functional group 'carbonyl' we designate aldehydic and ketone groups, but not

e.g. esters or carboxylic acids.".

Also at some other places in the text (e.g. the abstract) "carbonyls" is replaced by "aldehydes, ketones", or we make explicit that we refer to aldehydes and/or ketones.

"Oxo-acids", are these keto-acids only? Monocarboxylic acids and diacids are also oxo-acids in the general sense.

With oxo-acids, we meant both aldehydic and keto-acids. While there are no aldehydic acids present in our database, there could be in the future.

We have replaced "oxo-acids" by "carbonyl acids". A note in the table makes clear that we mean only aldehydic acids and keto-acids.

• p. 13273, Fig. 1: Are the data points of Ribeiro (1999, 2001)(s) really sublimation vapor pressures? This figure could be improved by using more different symbols and more emphasis on the symbol visibility (bolder) than on the lines (which are only intended to relate the different data points of the same reference).

From the description of Ribeiro (1999, 2001) it is clear that these are indeed sublimation pressures from crystalline samples. Perhaps the confusion arised because from 5 to 9 carbon atoms, no odd-even alternation is visible for these data. This is however due to the fact that Ribeiro (1999, 2001) provides no data on adipic (6 C) and suberic (8 C) atoms. By making the symbols bolder as suggested by the referee this will hopefully be more clear.

• p. 13274, Fig. 2: This figure should be improved concerning the horizontal positioning of the y-axis and the thickness of the frame borders.

The axis is shifted, the horizontal axis is changed to $\log_{10} \frac{p^0}{\text{atm}}$ and the vertical one to $\Delta H_v/(\text{kJ mol}^{-1})$. Note that three points from Yaws (1994) (citric acid, tartaric acid, malic acid) are added, confirming the correlation. Also one large outlier (falling beyond the scale of the original plot) from data of Monster (2004) is now included in the plot.

<center>C7473</center>

Last sentence in caption: "converted to subcooled liquid if necessary", better "converted to the subcooled liquid state, if necessary".

This has been done.

• p. 13275, Fig. 3: It is not mentioned in the figure caption nor the key, what the black line represents (presumably the 1:1 line).

"The black line is the 1:1 line." is added for Figs. 3, 4 and 5.

Also, it should be mentioned for clarity what the abbreviations stand for; better: "Left: modeled (est) vs. experimental (exp) log10 p0 for...".

This has been done.

I would also recommend the authors to mention in the figure caption that these plots represent the poor performance of the additive group-contribution model when applied to compounds >monofunctional, i.e. to compounds it was not optimized for.

We added: "This model, optimised for hydrocarbons and monofunctional compounds, overestimates the vapour pressure."

In this context, it would also be illustrative to show the improvement when the full method is used instead. Same issues apply to Fig. 4.

We have now also put the results of the full method here.

• p. 13277, Fig. 5: Since the dark blue and black symbols are hard to distinguish, different symbols would be helpful.

The dark blue is changed to light blue to improve contrast.

• p. 13278, Fig. 6: Caption: "with data points selected between", better: "with experimental data points selected between"

This has been done.

<center>C7474</center>

Technical Corrections

• Title: *"vapor pressure" is written using American English, while throughout the text the British English spelling "vapour pressure" is used.*

Replaced by "vapour pressure".

• p. 13230, l. 16: *"regulating its distribution", better: "influencing its distribution"*

Replacement done.

• p. 13230, l. 25: *"volatile organic compounds"*

This has been done.

• p. 13231, l. 4: *"formation Model Capouet et al., 2008;", spelling: "by Capouet et al. (2008)" or change to appropriate citation style with ().*

Style changed to "(Capouet et al., 2008)".

• p. 13231, l. 4: *", the GECKO-A", better: ", or the GECKO-A"*

This has been done.

• p. 13231, l. 5: *spelling: "Kinefics", same line correct citation style for (Aumont et al., 2005)*

Corrected.

• p. 13231, l. 6: *"leading to the oxygenated" delete "the"*

Done.

• p. 13231, l. 7: *"To simulate SOA formation, the chemical mechanism", to what is "the" referring to?*

To make the sentence clearer, we wrote: "To simulate SOA formation, such a chemical

C7475

mechanism can be coupled to"

• p. 13231, l. 9: *"partition to the SOA according to their vapour pressure" better: "partition to the condensed phase as a function of their vapour pressure" (to make clear that gas-particle partitioning is not governed by vapour pressure only).*

Done.

• p. 13231, l. 10: *"where the SOA is considered", better: "where the (organic) aerosol is considered" (whether the organic aerosol is secondary or primary is not of importance here).*

"SOA" replaced by "organic aerosol".

• p. 13231, l. 11: *check year of Cappa and Wilson citation*

Year changed to 2011.

• p. 13231, l. 28: *"assume additivity in ln p0", define here or above in the Introduction text the meaning of symbol p0.*

p^0 defined in the first sentence of the Introduction.

• p. 13231, l. 29: *"for all functional groups", maybe better: "with respect to contributions from different functional groups".*

Replacement done.

• p. 13232, l. 4: *"group-group combinations", maybe better: "group-group interactions"*

Replacement done.

• p. 13232, l. 8: *"very recently", delete "very"*

Deleted.

• p. 13232, l. 12: *"For these reasons, the need of a new estimation method addressing the above issues is clear.", better "For these reasons, a new estimation method*

C7476

addressing the issues mentioned above is desirable."

Sentence changed.

• *p. 13233, l. 2: "KDB" this abbreviation is not defined at this point.*

Definition of KDB is placed here.

• *p. 13233, l. 6: "the fusion temperature and enthalpy", maybe better: "the melting point temperature and enthalpy of fusion (see Section 2.2)"*

Sentence changed.

• *p. 13234, l. 26: "or or", spelling*

One "or" removed.

• *p. 13236, l. 1: no new paragraph*

Corrected.

• *p. 13239, l. 10: "this also applies to Ribeiro da Silva et al.", better: "this also applies to the data of Ribeiro da Silva et al."*

Corrected.

• *p. 13243, l. 18: delete "into"*

Deleted.

• *p. 13246, l. 27: "as they are counted already for k = 2." better: "as they are already accounted for by descriptor k = 2."*

Change inserted.

• *p. 13247, l. 11: "have lower vapour pressures than secondary alcohols", maybe write "corresponding secondary alcohols"*

"corresponding" inserted.

• *p. 13253, l. 25: replace "hitherto by us" by "in our previous SOA modelling studies"*

Replacement done.

Answer for referee #2

General questions

Do the authors believe the current implementation of EVAPORARTION would improve greatly as new data becomes available?

We have now gathered some extra data on mono acids, hydroxy acids and function-alised diacids. The extra data on mono acids (from 38 to 49 compounds) does not change significantly the parameters and the errors stay largely the same. We do not expect that introducing new data for monofunctional compounds will improve performance importantly, except for these classes where experimental data is limited and/or of lower accurracy (peroxides, hydroperoxides, peroxy acyl nitrates, peracids).

For hydroxy acids, we have now data for 10 instead of 5 compounds. The errors are now larger for this molecule class. For functionalised diacids we see the reverse: after including the extra data for malic, tartaric and citric acid, the model performance for functionalised diacids is significantly better. Both cases are reminiscent of the fact that when the data number is low, the reported errors are more speculative. In general, model improvement, as well as more reliable error estimation, can be expected for these molecule classes where the data number is currently low. We added therefore the following sentence on page 13252, line 24:

"Note that for molecule classes with only few compounds, the reported errors are rather speculative.".

At the end of the conclusions we added

"Within the present framework of EVAPORATION, better performance can be reached if data availability increases for these molecule classes where it is currently limited: e.g. peroxy acids, and hydroxy nitrates."

This is a very broad question but I wonder if there is a limiting accuracy of any model framework such as the one presented here. Do we know, as a community, what an acceptable level of accuracy would be. I get the feeling that we are some way from knowing this and require sensitivity studies using the wide range of gas phase degradation mechanisms available, comparing predicted and measured composition and mass. One useful way of assessing this would be to provide bands on, for example, figure 5 for the range of vapour pressures expected to be of major importance?

The main reason why we want to know the vapour pressure of a compound is because we want to know if, and to which degree, this compound partitions to the aerosol phase. In analogy to Valorso et al. (2011), we indicate the region of $10^{-13} - 10^{-5}$ atm as vapour pressure relevant for partitioning in the atmosphere. Above, the compound will not partition to an appreciable degree to the aerosol, even for large aerosol mass, while below the compound is essentially non-volatile, even for small aerosol mass. In Fig. 5, the data points in the upper right corner outside the area indicated by a dashed line, correspond to these compounds that are almost completely volatile, while the data points in the lower left corner correspond to these compounds that are almost completely non-volatile.

Furthermore, in figure 5 we provide now lines representing the vapour pressure needed to cause a change of 0.2 (black line), and of 0.5 (green line) in the condensed fraction ξ of a species, and this for a scenario where the aerosol mass concentration $C_{\mathrm{aer}} = 3.16$ μg/m^{-3}, $T = 298$ K, the mean aerosol molar mass $\overline{M}_{\mathrm{aer}} = 200$g mol^{-1} and ideality is assumed. An error of 0.2 on ξ can be considered significant, and an error of 0.5 grave.

From Fig. 5 it is clear that, certainly for vapour pressures close to $p* = \frac{C_{\mathrm{aer}} RT}{\overline{M}_{\mathrm{aer}}}$, rel-

atively small errors in p^0 result in important errors on ξ. To get the error on ξ below 0.2 at $p*$, the error on $\log_{10} p^0$ should be lower than 0.37. While for most classes of monofunctional and bifunctional compounds, the pred. MAD is below this value, this is not the case for most compound classes with more than two functional groups. To get the error on $\log_{10} p^0$ below 0.37 even for these compounds seems a major challenge. This is now also noted in the text.

Data weighting. In section 2.3 the authors specify the weighting factor given to experimental data vapour pressures. It is a little unclear whether 'one reference' refers to a specific source (e.g. Booth et al 2010) or a specific molecule (e.g. succinic acid). If it is the former, rather than the latter, how does one prescribe confidence in a specific set of estimations, such as multifunctional compounds, if there tends to be only one source? It might be my understanding but this is a little clear. I would presume the weighting factor takes into account the dominance of data from un-atmospherically representative compounds here, and this is the point the authors wish to make.

We acknowledge that our phrasing was not very clear on this. We reword lines 19 to 21 on page 13235:

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

In Eq. (3), "3" is replaced by "η".

Also, the choice of weighting using a factor of (1/3) seems very arbitrary. Have the authors assessed the impact of changing this factor? It is unclear at this point.

We did for the full model some optimisation tests with other choices: $\eta = 1, 5, 10, \infty$.

The impact on the parameter values is minor: g_k (the group contributions at 298 K) change on average by 0.01-0.03, and always less then 0.1. Therefore, we include after Eq. (3):

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

It would help if the authors would give examples of vapour pressure estimations using their new model..for simple hydrocarbons and more complex molecules.

As noted to the first reviewer, we will add in the supplementary material a list of examples.

As with most fitted group methods, the true test arises from applying the model to data not used within the fitting procedure. This is partly why we know the existing models do not perform very well. How do the authors plan on testing this new method beyond the safe confinds of the data used within the fitting procedure? Can the authors comment on if there is any advantage in refitting existing frameworks despite some of the failings in the frameworks (such as group location).

With the use of the leave-one-out method we already applied a cross-validation procedure. A more stringent test can be done if sufficiently new data becomes available. Then it will become clear if (i) EVAPORATION can be kept in its current form, (ii) a reparameterisation is necessary or (iii) a change of the framework.

In our opinion, the main advantage of refitting an existing framework is that it involves less work than devising a new framework.

Minor points:

Please check the reference to abbreviations throughout. It can be quite hard to remember what each refers to. I would suggest re-iterating some definitions.

Table 8 and section 5 already gives an overview of the abbreviations for vapour pressure methods. We have added references at a few places, but we don't think re-iterating definitions would improve the style of the manuscript.

Page 13235: Equation 1.. reference is missing.

Reference inserted (see answer for referee # 1).

Page 13244. Where has equation 22 come from?

This is actually a definition. So we rephrase: "We report also the total group contribution g_k at 298 K, defined as"

Page 13245, line 12. I would suggest another reference other than Wikipedia is used.

We have inserted instead as definition

" 'the number of independent units of information in a sample relevant to the estimation of a parameter' (Everett, 2010)."

with the reference Everett, B. S., The Cambridge Dictionary of Statistics, Cambridge University Press, Cambridge, UK, 2nd edn., 2002.

Answer to referee # 3

The distinction between the "simple" additive group contribution method and the "full" model is not completely clear at several locations in the text as well as in the figures. In particular it is not clear to me whether in Fig. 3 and 4 the agreement between modeled

and experimental data would improve if the full model of Table 4 would be used instead of the additive model of Table 2.

See specific comment of referee #1. The estimations of the full model have been added. The agreement is better with the full model.

It would help the reader to understand the different descriptors better if for a few example compounds their assignment would be illustrated (perhaps in an appendix)

See comment referee #1. We have included supplementary material with examples.

Minor comments:

page 13234, line 23: Booth et al., 2011 should be added to the references.

This is done.

page 13250, first paragraph: While I agree that the differences between a "simple" group contribution method and the measurements of Booth et al. (2010) can not be explained by uncertainties between different types of measurements, there is the possibility that there are problems associated with the measurements of the vapor pressures of the solids as discussed in Soonsin et al. (2010). If you compare the sub-cooled vapor pressures derived from measured vapor pressures of the solids for the dicarboxylic acids of Booth et al. (2010) with direct measurements of the sub-cooled liquids (e.g. Soonsin et al., 2010, Riipinen et al., 2007) they agree within error for succinic and glutaric acid while they disagree for malonic acid.

The discrepancies between vapour pressure measurements for solids (or saturated solutions) of different research groups is indeed obvious. From malonic to glutaric acid, the sublimation pressures of Soonsin (2010) are about a factor 10.8, 7.5 and 7.5 lower than those of Booth (2010). There are problems with at least one set of measurements for the solid state.

The reviewer presents the comparison of liquid vapour pressures (Soonsin, 2010, Riipinen, 2007), and solid vapour pressures converted to subcooled liquid by taking fusion data into account (Booth, 2010). This comparison actually entails considerable complexity. The data of Soonsin (2010) and Riipinen (2007) use an activity coefficient model to get the vapour pressure of the pure liquid. Koponen et al. (2007) showed that for malonic acid, there can be large differences when using different activity coefficient models (a factor 39), while for succinic acid and glutaric acid, the variation was much smaller (about a factor 1.3).

However, the argument of uncertainty in activity coefficient model applies much less for Soonsin (2010), as they validated their activity coefficient model with measured activities over a wide range of mixture composition for malonic, succinic and glutaric acid. This gives some confidence in the liquid data of Soonsin (2010), and, as they agree, also in the liquid data of Riipinen (2007).

The reviewer states that the sub-cooled vapour pressures of succinic and glutaric acid of Booth (2010), derived from solid vapour pressures by taking fusion data into account, agree within error with the liquid vapour pressure data of Soonsin (2010) and Riipinen (2007). On the other hand, Soonsin *et al.* (2010) state that their solid and liquid vapour pressure data agree with the fusion data of these acids. This is at first sight a contradiction, as the solid state vapour pressure data of Soonsin (2010) is much lower than that of Booth (2010).

The contradiction is resolved as one recognizes that two different extreme assumptions were used in the solid-to-liquid conversion. Booth (2010) assume $\Delta C_{p,ls} = \Delta S_{\text{fus}}$, while Soonsin (2010) assume $\Delta C_{p,ls} = 0$. With this assumption, the data of Booth (2010) is a factor of 10, 5.5 and 2.5 higher for malonic, succinic and glutaric acid respectively (with the assumption of $\Delta C_{p,ls} = \Delta S_{\text{fus}}$ it was a factor of 7.4, 2.3 and 2.1). The assumption of $\Delta C_{p,ls} = 0$ would then give disagreement between Soonsin (2010) and Booth (2010) not only for malonic acid but also for succinic acid.

The important point here is that the lack of knowledge on $\Delta C_{p,ls}$ brings considerable uncertainty for these compounds with high T_{fus} and ΔH_{fus}, if a solid state vapour pressure is converted to a subcooled one.

Given the complexity of comparing liquid vapour pressures with solid vapour pressures converted to subcooled liquid, we do not give this example in the text but confine ourselves to the comparison of sublimation pressures. Therefore we added:

"A possible explanation could be that there are problems with the measurements of sublimation pressures of functionalised diacids, as Soonsin et al. (2010) had already concluded on the sublimation pressures of diacids. Unfortunately, as long as there are no other room-temperature sublimation pressure measurements available, this is difficult to verify."

It would be interesting to see measurements of tartaric and citric acid in their liquid state. This discussion is given in the section conclusion, but it would be helpful for the reader if it is mentioned also at this point.

We have now found data of malic, tartaric and citric acid in their liquid state, albeit of a secondary source (Yaws), and at high temperature. According to these high temperature data, tartaric acid has a vapour pressure lower than succinic acid, and citric acid lower than adipic acid. This is more in agreement with the intuitive idea that adding more functional groups lowers the vapour pressure. Moreover, their (boldly extrapolated) vapour pressures fall nicely on the $\Delta H_v \left(298\,\mathrm{K}\right) - \log_{10} p^0 \left(298\,\mathrm{K}\right)$ correlation, as was already the case for the high-temperature linear diacid data from ESDU. More insight can be provided if room-temperature subcooled liquid vapour pressure data for citric and tartaric acid becomes available, analogous to the measurements of Soonsin et al. (2010) on diacids.

Therefore, we add at line 14 of page 13250:

"Furthermore, we note that according to the high-temperature (above the melting point)

liquid vapour pressure data of Yaws (1994), citric and tartaric acid have a lower vapour pressure than respectively adipic acid and succinic acid, more according to chemical intuition. "

page 13254, line 14: write instead of MG: Marrero and Gani (2001)

This has been done.

Figures: I find Fig.6 very difficult to read, I suggest to print left and right in either separate figures or over two columns so that the assignment to each compound class is easier to see.

The figure spans now two columns.

Other changes

(1) In Table 4, I added some details on how to count the group interaction effects.

(2) The insertion of the high-temperature liquid vapour pressure data on functionalised diacids from Yaws (1994) necessitates a major change in the conclusions, page 13256, line 4-16:

"Counter-intuitively, the subcooled liquid data -calculated from sublimation pressures- of Booth et al. (2010) suggest that citric and tartaric acid have higher vapour pressures than adipic and succinic acid respectively although they have more polar groups. The high-temperature (above the melting point) liquid vapour pressure data of Yaws (1994) however suggest the reverse. Moreover, the data of Yaws (1994) on functionalised diacids, and of ESDU on diacids, after bold extrapolation, does obey the ΔH_v vs. $\log_{10}(p^0)$ correlation at 298 K already established for a wide variety of compounds (MacLeod et al., 2007, Epstein et al., 2010). Most ambient temperature data on diacids and functionalised diacids does not (except e.g. Cappa et al. 2007). One possible ex-

planation is that there are problems with the measurements. More light can hopefully be shed on this issue by the measurement of high-temperature (above the melting point) liquid vapour pressure of other functionalised diacids. Also confirmation of the data of Yaws (1994) is desired, as this is a secondary source with no details on the experimental procedure. The high-temperature liquid vapour pressures should be relatively reliable: no solid to subcooled liquid (e.g., Booth et al., 2010) or mixture to pure liquid conversion (e.g., Riipinen et al., 2007) would be needed, and the vapour pressure should be more accurately measurable at this higher temperature. Also room temperature measurement of subcooled liquid vapour pressure of functionalised diacids by the methodology of Soonsin et al. (2010) can provide more insight, by comparing it with the existing solid vapour pressure data."

References

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.

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

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 13229, 2011.

C7487