Response to Reviewers

We sincerely thank the Editor and the reviewers for their valuable time and insightful comments on our manuscript. We appreciate the constructive feedback, which has helped us significantly improve the quality and clarity of the paper. All comments have been carefully addressed, and the manuscript has been revised accordingly. Below, we provide a point-by-point response to each comment raised by the reviewers. All changes in the manuscript are highlighted in yellow color.

Reviewer 1

Bezaatpour et al. (2025) have provided a very valuable paper, and associated tool, to the environmental science community. The method and validation is very comprehensive, and the paper is thoughtful in exploring the limitations of the presented tool and its implemented estimation method. I recommend publication following the minor revisions suggested below.

Introduction

Comment: Please state somewhere the type of saturation vapour pressure being estimated, including the Raoult effect and the phase, e.g.: 'pure component liquid (or sub-cooled liquid) saturation vapour pressure'. (where pure component deals with the Raoult effect and liquid (or sub-cooled liquid) deals with the phase.

Response: *In the revised manuscript, we stated that the model calculates the pure-component liquid (or sub-cooled liquid) saturation vapor pressure, which corresponds to the vapor pressure of a compound in its pure state (without accounting for Raoult's law mixing effects) and in the liquid or sub-cooled liquid phase, as follows:*

In its current version, VaPOrS implements the SIMPOL method by detecting 30 functional groups required for estimating pure-component (sub-cooled) liquid saturation vapor pressure and the corresponding enthalpy of vaporization, without accounting for Raoult's law effects.

Comment: Line 6 of Page 5, could you reference the relevant results section of your paper when discussing the UManSysProp inaccuracies? For readers interested in this inaccuracy, doing so will save them time searching for related information.

Response: *In the revised manuscript, we added a reference to the relevant section (i.e., 4.3) discussing the UManSysProp inaccuracies.*

Section 3

Comment: Line 8 Page 30 and Line 14 Page 31 and Line 14 Page 33

Whereas the authors reference the original SIMPOL paper for the source of measured vapour pressures in Fig. 4. It's unclear where the measurements for Fig. 5, Fig. 6 and Fig. 7 come from, please provide reference(s).

Response: *The measured vapor pressures used for validation in Figures 5–7 were obtained from the same experimental datasets reported in the original SIMPOL study by Pankow and Asher (2008). In the revised manuscript, we clarified the sources of the measured vapor pressure data. The corresponding reference has now been explicitly cited in the figure captions and in the text.*

Results

Comment: Line 7 Page 36

In terms of what is available to users from the internet, the current MCM website (https://www.mcm.york.ac.uk/MCM/ accessed October 15th 2025) does not provide a file containing SMILES of all MCM species. This has been confirmed with Killian Murphy and Andrew Rickard at University of York. If it helps the authors direct readers to such a file, for the current version of MCM, v3.3.1, a file is stored here, with no plans to remove it: https://github.com/simonom/PyCHAM/blob/master/PyCHAM/prop_store/mcm_v3p3p1_all_species.xml.

I see (code availability section) that the MCM v3.3.1 SMILES may also be available in the zenodo repository. If so, please state explicitly in the code availability whether this is the case. Either way, please make clearer at line 7 page 36 where readers can obtain the MCM v3.3.1 SMILES.

Response: *In the revised manuscript, we stated that the current version of the MCM (v3.3.1) can be obtained as the file mcm_3-3-1_species_complete.tsv from the MCM archive page (https://www.mcm.york.ac.uk/MCM/about/archive)*

Code availability

Comment: Could a URL or DOI be provided for the relevant zenodo repository?

Response: *In the revised manuscript, we added the following DOI to the Code availability section: https://doi.org/10.5281/ZENODO.15222175*

Reviewer 2

The manuscript has improved substantially and is now ready for publication. I'm pleased to see that the main points raised by the reviewers have been addressed. The focus has been shifted toward highlighting the novelty and versatility of the group recognition framework, with the SIMPOL implementation presented as a case study demonstrating its potential applications. The addition of the new tables provides validation data that illustrate the framework's superior performance compared to similar tools.

In my opinion, the fact that the tool does not perform SMILES standardization and relies on the user to provide correct input SMILES remains a limitation. However, this limitation is transparently acknowledged in the manuscript. I'm glad to see that the authors have plan for address this in the future.

Comment: I think the conclusions of the paper feel heavily written by AI (most of the acronyms are spelled out even if were explained before, please address this). I'd rather a more "human-like" tone.

Response: *We appreciate the reviewer's comment. Initially, we had expanded acronyms in the Conclusions section to ensure clarity for readers who might focus only on this section. However, based on the reviewer's suggestion, we have now removed these expansions to make the conclusions read more naturally while retaining all key information.*

Comment: Finally, on page 7, line 15, please remove the word "unique." Canonical SMILES do not represent a unique molecular representation.

Response: *We thank the reviewer for this comment. We agree that the term "unique" is not accurate in this context, as canonical SMILES do not guarantee a single molecular representation. Accordingly, we have removed the word "unique" from Page 7, Line 15 in the revised manuscript.*

Reviewer 3

Comment: wherever an estimated vapour pressure is provided, the corresponding temperature should be stated in the main text and Figure/Table caption, for example the Figures in Section 4.2 and Tables in Section 4.3. Further clarity is needed when vapour pressures from 'SIMPOL' or the 'SIMPOL method' are provided. If these have been calculated manually, this should be stated in each sub-section of results, including in Sections 3.1 and 4.3. Or, if these have been obtained from the original Pankow and Asher paper (e.g. using the Antoine coefficients in their supplement), this should be stated (including where in the paper they come from). If there is an alternative source, then this should be stated. Otherwise the assumption is that values have come directly from the original paper, which could mean a different interpretation than for manual generation.

Response: *We thank the reviewer for this suggestion. In the revised manuscript, all estimated vapor pressures now explicitly state the corresponding temperature in the main text as well as in figure and table. Furthermore, we clarified the source of all vapor pressures labeled as "SIMPOL", "VaPOrS", "Experimental" and "UManSysProp". These updates ensure readers can distinguish between calculated and literature-derived values throughout the relevant sections.*

Comment: There is a degree of circularity that arises from the same authors writing the matching patterns for the VaPOrS tool and manually estimating vapour pressures, when both are based on their interpretation of the SIMPOL rules. Ideally the VaPOrS estimates would be compared against independent estimates, e.g. from the original SIMPOL paper, though unfortunately that paper does not seem to supply the estimates in tabulated form. But there is some interpretation of the SIMPOL rules needed, and some of the chemicals presented in Section 4.3 exemplify the point. E.g. dimethyl-hydroxylamine is strictly a hydroxylamine, not an amine, though my calculations show that the authors have included the secondary amine group to achieve the manually estimated (SIMPOL) vapour pressure. And to be fair, Table 1C of the Pankow and Asher paper lists this molecule as an amine, suggesting a consistency with the intended meaning of this group. The same molecule includes a hydroxyl group bonded to a nitrogen atom, but both Table 5 of the Pankow and Asher paper and Section 2.1.5 of the paper under review refer to the hydroxyl group to be considered for the group contribution as an alkyl hydroxyl, which my understanding means the hydroxyl group bonded to a non-aromatic carbon atom. Arguably

because the bond is to nitrogen rather than carbon the hydroxyl group should not contribute to the estimated vapour pressure in this case, but my calculations show that the authors have included it for their manual SIMPOL estimate. I think the authors need to acknowledge at least these points of interpretation in the paper, plus any other points where interpretation has been needed for the SIMPOL rules.

Response: *We appreciate the reviewer's insightful observation regarding the interpretation of the SIMPOL functional group definitions, particularly in relation to compounds such as dimethyl-hydroxylamine. We acknowledge that certain group classifications required interpretative judgment since the original SIMPOL paper did not provide tabulated vapor pressure values or exhaustive group assignments for all compounds. Our classification choices were guided strictly by the published definitions and data representations in Pankow and Asher (2008). Specifically, while the SIMPOL framework defines alkyl hydroxyls as hydroxyl groups bonded to non-aromatic carbon atoms, our observation of Figures 7(a) and 13 indicated that the hydroxyl group in dimethyl-hydroxylamine contributed to both vapor pressure and enthalpy of vaporization. Since the numerical data were not available despite our direct request from the original authors, our interpretation was necessarily based on visual analysis of the published figures. We therefore adopted this interpretation in VaPOrS to remain consistent with the apparent treatment in the SIMPOL study. Importantly, VaPOrS is designed to be flexible and can be updated to incorporate any future clarifications or refinements provided by the SIMPOL authors or by users of the tool. We addressed this matter in the revised manuscript as follows:*

It is worth mentioning that in compounds such as dimethyl-hydroxylamine, the initial classification considered only the amine functional group, excluding the hydroxyl, as the SIMPOL definition specifies alkyl hydroxyls as hydroxyl groups bonded to non-aromatic carbon atoms. However, careful observation of Figures 7(a) and 13 in (Pankow and Asher 2008) revealed that the hydroxyl group appears to have been included in the calculation of both vapor pressure and enthalpy of vaporization. Although the numerical values were unavailable, this observation led us to infer that the phrase "non-aromatic carbon" may have been used to distinguish the behavior of hydroxyl groups attached to aromatic rings rather than to strictly define the atom of attachment. Accordingly, VaPOrS adopts this interpretation to align with the apparent implementation in SIMPOL. The tool remains adaptable and can be updated should new clarifications or refinements to the SIMPOL framework become available.