

Comment

Comment on Naef, R.; Acree, W.E., Jr. Calculation of the Three Partition Coefficients logPow, logKoa and logKaw of Organic Molecules at Standard Conditions at Once by Means of a Generally Applicable Group Additivity Method. *Preprints* 2023, 2023120275

Robert J. Meier 

Independent Researcher, 52525 Heinsberg, Germany; r.meier@planet.nl

Next to the paper referred to in the title [1], Naef and Acree have published a series of papers involving the group contribution (GC) method, which has been applied to a larger variety of physico-chemical properties all published in *Molecules* [2–8]. The GC method is based on the assumption that some aspects of chemical groups are the same in many different molecules. The GC method is a so-called data-driven model, with experimental data being used to parametrize the model. A molecular property of interest is the sum (additive) of the individual properties of molecular fragments j , with N_j being the number of times the fragment j occurs in the molecule, i.e.,

$$\Delta H_f = \sum_{j=1,N} N_j \cdot \Delta H_f(j) \quad (1)$$

We talk more specifically about the group additivity method, as in the title of the paper by Naef and Acree.

Regarding the former paper by Naef and Acree on the heat of formation of organic molecules [2], we criticized this publication as the authors had grossly overlooked very many relevant publications over the last 25 years [9]. There have been numerous papers by several excellent groups that reported significant progress on this topic with partly excellent results. Interestingly, Naef and Acree did not provide a rebuttal to the Comment [9]. We were more astonished when seeing the latest contribution by Naef and Acree [1], which addresses the octanol–water partition coefficient log Kow. The prediction of log Kow, or alternatively log P, has been of considerable interest in various fields and was therefore often studied; a huge number of reports as well as software tools are available. Similarly to the case of the heat of formation, Naef and Acree state now that their results ‘clearly outperform Klopman’s results’; however, Klopman’s paper dates back to 1994! Apart from many individual papers, also-important review papers [10,11] are simply not quoted, and do not discuss comparisons of their results to the state of the art, which should be key in any scientific publication. Furthermore, there is ample literature not quoted by Naef and Acree, and there is a larger variety of software tools both freely available as well as commercial ones (a few are mentioned by Naef and Acree, but no proper analysis is provided). In 2002, Marrero and Gani reported a GC-based study on log Kow [12] with the result, ‘The group-contribution values were calculated by linear regression analysis using a data set of 9560 values for K_{ow} . The data set included compounds ranging from C3 to C70, including large and heterocyclic compounds. Compared to other currently used group-contribution methods, the new methods make significant improvements in accuracy with logarithm-unit average absolute errors of 0.24 for K_{ow} ’. In comparison, Naef and Acree reported results based on 3332 molecules (about 1/3 of the Marrero–Gani work) with a standard deviation of 0.42. Firstly, it needs to be mentioned that the earlier value from Naef and Acree was



Citation: Meier, R.J. Comment on Naef, R.; Acree, W.E., Jr. Calculation of the Three Partition Coefficients logPow, logKoa and logKaw of Organic Molecules at Standard Conditions at Once by Means of a Generally Applicable Group Additivity Method. *Preprints* 2023, 2023120275. *Molecules* 2024, 29, 892. <https://doi.org/10.3390/molecules29040892>

Academic Editor: Antonio Fernández-Ramos

Received: 28 December 2023

Revised: 2 February 2024

Accepted: 15 February 2024

Published: 18 February 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

0.51, whereas the lower value of 0.42 is the result of the removal of 122 molecules (3.5% of the total set) from the parameter computation for which the experimental value deviates by more than three times the value of *S*. As the *S*(tandard) *D*(eviation) is more sensitive to outliers, the results of Marrero–Gani and Naef and Acree might still be close; however, the latter is not better, whereas the former result dates back to 2002. The Marrero–Gani work was based on three times as many species, which implies a larger range of validation and an absolute average error, which compares well with the experimental error that is typically in the range of 0.2–0.4 log units [13]. At a later stage, the results from the same group reported a standard deviation of 0.61 and an averaged absolute deviation of 0.45 log units based on 12,193 different molecules. The results and many more properties were reported in the thesis of Amol Shivajirao Hukkerikar (2013), which is freely available [14], as well as in a paper [15]. We wish to emphasize that only looking at the standard deviation or a similar quantity is insufficient; the number of different molecules is crucial regarding the proven applicability range.

A more recent paper considers log *K*_{ow} with deep learning techniques [13]. In addition, interestingly, the software tool developed has been made available via the Supplementary Material. Different DNN models were explored based on a data set comprising 14,050 chemicals. The authors have considered both neutral as well as ionic species. The data set itself was said to be heterogeneous, including many different classes of chemical compounds. After the exclusion of some identified erroneous data points, a root mean square error (rmse) of 0.47 and a very small standard deviation was reported; for details, refer to Ref. [13].

There are more works on log *K*_{ow} that are of relevance. In summary, it seems justified to state that whereas Naef and Acree claim that their method ‘has proven its unmatched versatility in the equally reliable prediction of up to now 19 physical, thermodynamic, solubility-, optics-, charge-, and environment-related molecular descriptors based on a common group-additivity method’, this is in such a general way that has not been supported by the literature of the last 20 years. The statement ‘enabled their prediction for nearly 29,500 molecules’ merely means that the tool developed based on much fewer species is simply applied to 29,500 molecules, whereas other existing methods have performed, as proven, with respect to that many species.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Naef, R.; Acree, W.E., Jr. Calculation of the Three Partition Coefficients log*P*_{ow}, log*K*_{oa} and log*K*_{aw} of Organic Molecules at Standard Conditions at Once by Means of a Generally Applicable Group-Additivity Method. *Preprints* **2023**, 2023120275. [CrossRef]
2. Naef, R.; Acree, W.E., Jr. Revision and Extension of a Generally Applicable Group-Additivity Method for the Calculation of the Standard Heat of Combustion and Formation of Organic Molecules. *Molecules* **2021**, *26*, 6101. [CrossRef] [PubMed]
3. Naef, R. A Generally Applicable Computer Algorithm Based on the Group Additivity Method for the Calculation of Seven Molecular Descriptors: Heat of Combustion, Log*P*_o/*W*, Log*S*, Refractivity, Polarizability, Toxicity and Log*B*_B of Organic Compounds; Scope and Limits of Applicability. *Molecules* **2015**, *20*, 18279–18351. [PubMed]
4. Naef, R.; Acree, W.E. Calculation of Five Thermodynamic Molecular Descriptors by Means of a General Computer Algorithm Based on the Group-Additivity Method: Standard Enthalpies of Vaporization, Sublimation and Solvation, and Entropy of Fusion of Ordinary Organic Molecules and Total Phase-Change Entropy of Liquid Crystals. *Molecules* **2017**, *22*, 1059. [PubMed]
5. Naef, R.; Acree, W.E. Application of a General Computer Algorithm Based on the Group-Additivity Method for the Calculation of Two Molecular Descriptors at Both Ends of Dilution: Liquid Viscosity and Activity Coefficient in Water at Infinite Dilution. *Molecules* **2018**, *23*, 5. [CrossRef] [PubMed]
6. Naef, R.; Acree, W.E., Jr. Calculation of the Surface Tension of Ordinary Organic and Ionic Liquids by Means of a Generally Applicable Computer Algorithm Based on the Group-Additivity Method. *Molecules* **2018**, *23*, 1224. [CrossRef] [PubMed]
7. Naef, R. Calculation of the Isobaric Heat Capacities of the Liquid and Solid Phase of Organic Compounds at 298.15 K by Means of the Group-Additivity Method. *Molecules* **2020**, *25*, 1147. [CrossRef] [PubMed]
8. Naef, R.; Acree, W.E., Jr. Calculation of the Vapour Pressure of Organic Molecules by Means of a Group-Additivity Method and their Resultant Gibbs Free Energy and Entropy of Vaporization at 298.15 K. *Molecules* **2021**, *26*, 1045. [CrossRef] [PubMed]

9. Meier, R.J. Comment on Naef, R.; Acree, W.E., Jr. Revision and Extension of a Generally Applicable Group-Additivity Method for the Calculation of the Standard Heat of Combustion and Formation of Organic Molecules. *Molecules* 2021, 26, 6101. *Molecules* 2023, 28, 6215. [[CrossRef](#)]
10. Katritzky, A.R.; Kuanar, M.; Slavov, S.; Hall, C.D.; Karelson, M.; Kahn, I.; Dobchev, D.A. Quantitative correlation of physical and chemical properties with chemical structure: Utility for prediction. *Chem. Rev.* 2010, 110, 5714–5789. [[CrossRef](#)] [[PubMed](#)]
11. Nieto-Draghi, C.; Fayet, G.; Creton, B.; Rozanska, X.; Rotureau, P.; De Hemptinne, J.-C.; Ungerer, P.; Rousseau, B.; Adamo, C. A General Guidebook for the Theoretical Prediction of Physicochemical Properties of Chemicals for Regulatory Purposes. *Chem. Rev.* 2015, 23, 13093–13164. [[CrossRef](#)] [[PubMed](#)]
12. Marrero, J.; Gani, R. Group-Contribution-Based Estimation of Octanol/Water Partition Coefficient and Aqueous Solubility. *Ind. Eng. Chem. Res.* 2002, 41, 6623–6633. [[CrossRef](#)]
13. Ulrich, N.; Goss, K.-U.; Ebert, A. Exploring the octanol–water partition coefficient dataset using deep learning techniques and data augmentation. *Commun. Chem.* 2021, 4, 90. [[CrossRef](#)] [[PubMed](#)]
14. Hukkerikar, A.S. Development of Pure Component Property Models for Chemical Product-Process Design and Analysis. Ph.D. Thesis, Danish Technical University, Kongens Lyngby, Denmark, 2013. Available online: https://backend.orbit.dtu.dk/ws/portalfiles/portal/59650003/Amol+S.+Hukkerikar_PEC13-42.pdf (accessed on 6 March 2023).
15. Hukkerikar, A.S.; Sarup, B.; TenKate, A.; Abildskov, J.; Sin, G.; Gani, R. Group-contribution⁺ (GC⁺) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilibria* 2012, 321, 25–43. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.