Accurate prediction of standard enthalpy of formation based on semiempirical quantum chemistry methods with artificial neural network and molecular descriptors

zhongyu Wan[1], Quan-de Wang[2], and Jinhu Liang[3]

[1]Xuzhou Institute of Technology
[2]China University of Mining and Technology
[3]North University of China

June 5, 2020

Abstract

This work investigates possible improvements in the accuracy of semiempirical quantum chemistry (SQC) methods for the prediction of standard enthalpy of formation ($\Delta_f H^o$) through the use of artificial neural network (ANN) with molecular descriptors. A total number of 142 organic compounds with enough structural diversity has been considered in the training set. Standard enthalpy of formation for the selected compounds at the semiempirical PM3 and PM6 quantum chemistry methods is collected from literature, and is calculated by using semiempirical PM7 method in this work. The multiple stepwise regression is first employed to screen effective molecular descriptors, which are highly correlated with the error terms of the standard enthalpy of formation compared with experimental values. The obtained 7 effective molecular descriptors are then used as input set to establish three 7-11-1 neural network-based correction models to improve the accuracy of SQC methods. By using the developed correction models, the mean absolute errors (MAE) for $\Delta_f H^o$ of PM3, PM6, and PM7 methods are reduced from 22.36, 18.60, 17.27 to 9.86, 9.83, 8.95, respectively in kJ/mol. Meanwhile, the results of the test set show that the neural network does not have the problem of over-fitting. Detailed analysis of the 7 effective molecular descriptors indicates that the major source to the correction models is from the electron withdrawing effect. The developed ANN models for the three selected SQC methods provide an efficient method for the quick and accurate prediction of thermodynamic properties.

Hosted file

`manuscript_IJQC.docx` available at https://authorea.com/users/330249/articles/457092-accurate-prediction-of-standard-enthalpy-of-formation-based-on-semiempirical-quantum-chemistry-methods-with-artificial-neural-network-and-molecular-descriptors