Electronic structure

https://doi.org/10.1038/s43588-022-00393-z

Using AI to navigate through the DFA zoo

Stefan Vuckovic

A proposed density functional approximation (DFA) recommender outperforms the use of a single functional by selecting the optimal exchange-correlation functional for a given system.

As the workhorse of quantum modeling, density functional theory (DFT) has become one of the largest consumers of the world's supercomputing power[1]. While in principle DFT is an exact theory, in practice, density functional approximations (DFAs) must be employed to make simulations feasible. Thus, whenever a DFT calculation is run to solve a given problem in chemistry, biology, or materials science, one needs to pick one of the hundred available DFAs from the 'DFA zoo'. As no currently available DFA from this zoo displays universal accuracy across chemistry, researchers have devoted much effort to find the most faithful DFA for a chemical problem at hand. In their article in *Nature Computational Science*, Duan et al. use artificial intelligence (AI) to navigate through the DFA zoo by building machine learning models that automatically pick the best DFA for a problem at hand without requiring human intervention or trial-end-error procedures[2].

DFT is nowadays routinely used to rationalize, guide, and support experiments. But the full potential of DFT for transforming chemical discovery will be unlocked only after it becomes fully predictive. The recent advances in combining AI with DFT enabled massive accelerations in automated chemical discoveries that use DFT as the foundation[3]. For example, DFT empowered by AI has been used to screen millions of compounds for specific applications[3]. However, whether such large-scale screenings will give us the 'right' lead compounds will ultimately depend on the accuracy of the employed DFA. The main issue is that DFAs are not universal, meaning that some work better for main-group molecules, some for transition metal complexes, and some for solids. The DFA zoo is constantly growing, making the problem of picking the right DFA even more difficult. One hope is that, in five to ten years from now, the DFT community (likely with the assistance of AI[4] and/or the new building blocks with a more sophisticated mathematical structure drawing from the exact theory[5]) will develop a DFA that displays far more universal accuracy than the ones present in the current zoo.

However, until the DFT methodology becomes more universal, we have to deal with the currently available zoo. Duan et al. provide a framework for better navigation through the DFA zoo by using an AI procedure for the automatized recommendation of what DFA should be applied to a problem at hand. The authors train their AI models to give the correction to DFT results for each selected DFA. Then, the recommended DFA will be the one for which the AI model predicts the smallest correction. This 'correction and recommender' strategy beats in accuracy any individual DFAs for extremely challenging energetics of transition metal complexes at different spin states. The basics behind their DFT recommender are illustrated in Fig. 1. Starting from a given transition metal complex of interest (Fig. 1a), one runs a series of DFT calculations by using a subset of the zoo's DFAs (Fig. 1d). The information extracted from the electronic density (that is, the probability measure of finding an electron at a given point in space) obtained from one of the DFT calculations (Fig. 1b) is then passed to a machine learning model (Fig. 1c), which, in turn, gives a correction to DFT results and a recommendation on what approximation should give the most accurate results for the given transition metal complex (Fig. 1d).

Duan et al. make important steps in using AI to navigate through the DFA zoo and improve DFT results. The resulting methodology can be coupled with the existing DFT/AI workflows for chemical discovery[3]. However, important challenges still remain. The proposed AI model can only be applied to specific types of energies and chemical elements that have been used in the training. Second, the quality of the training (reference) data needs further improvements, as even

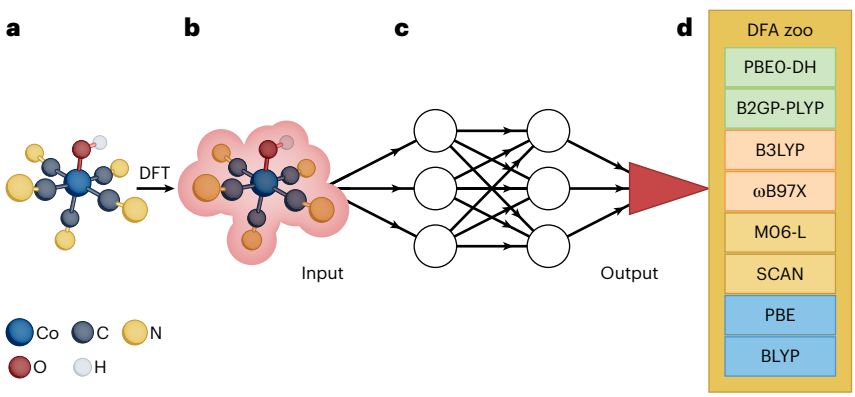

Fig. 1 | DFT recommender of Duan et al. a–c, For a given (transition metal) molecule (**a**), the electronic density from a DFT calculation (red shading in **b**) is used as an input to a machine learning model (**c**), which outputs the correction to the DFT results and makes a recommendation on what DFT approximation should give the most faithful results (red arrow). **d**, A selection of DFAs from the zoo used by Duan et al., where different colors represent different classes of DFAs. Blue, generalized gradient approximations (GGAs); yellow, meta-GGAs; orange, hybrid DFAs; green, double hybrids. The selected DFA (ωB97X) is picked here as the best just as an illustration. In panel **a**, a cobalt(II) transition metal complex has been used as an example.

the 'gold standard' methodology that the authors use to benchmark DFT and train their AI models does not always give faithful results for transition metal complexes[6], which are generally very challenging for quantum modeling. Nevertheless, the framework that the authors propose is general and can be further improved once more accurate benchmarks become available.

Stefan Vuckovic
Physical and Theoretical Chemistry, University of Wuppertal, Wuppertal, Germany.
e-mail: stefanvuckovic1@gmail.com

Published online: 26 January 2023

References

1. Kulik, H. J. et al. *Electron. Struct.* **4**, 023004 (2022).
2. Duan, C., Nandy, A., Meyer, R., Arunachalam, N. & Kulik, H. J. *Nat. Comput. Sci.* https://doi.org/10.1038/s43588-022-00384-0 (2022).
3. Nandy, A. et al. *Chem. Rev.* **121**, 9927–10000 (2021).
4. Kalita, B., Li, L., McCarty, R. J. & Burke, K. *Acc. Chem. Res.* **54**, 818–826 (2021).
5. Vuckovic, S. et al. *WIREs Comput. Mol. Sci.* https://doi.org/10.1002/wcms.1634 (2022).
6. Zhang, D. & Truhlar, D. G. *J. Chem. Theory Comput.* **16**, 4416–4428 (2020).

Competing interests

The author declares no competing interests.