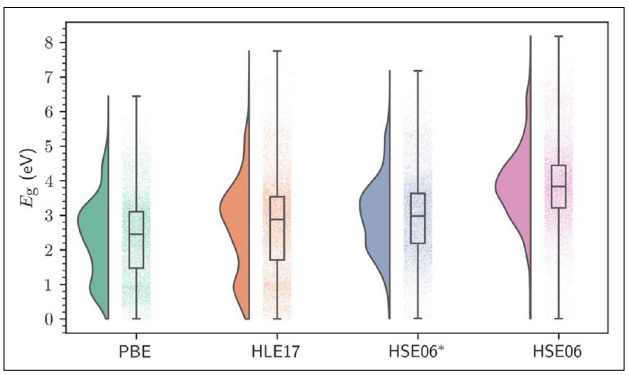
Machine learning improves metal–organic frameworks design and discovery

An impeding challenge that confronts materials scientists is accurately and rapidly predicting the electronic properties of metal–organic frameworks (MOFs). In a recent report published in *NPJ Computational Materials* (https://doi.org/10.1038/s41524-022-00796-6), researchers have leveraged the Quantum MOF (QMOF) Database that provides data sets of computed quantum chemical properties, including the electronic bandgap, of over 20,000 MOFs and coordination polymers.

Andrew Rosen, a Miller Research Fellow at the University of California, Berkeley (UC Berkeley), is the lead scientist of this work. He collaborated with a team of materials researchers, chemists, and chemical engineers from UC Berkeley, Lawrence Berkeley National Laboratory (LBNL), and Northwestern University. They discovered a robust methodology for accelerating accurate predictions of MOF bandgaps by using density functional theory (DFT) and machine learning. Rosen underlines the significance of the QMOF Database: "It can require thousands of computing hours to predict a MOF bandgap with density functional theory. [O]ur machine learning model reduces the compute time to less than a second per material with high accuracy."

A MOF is a hybrid class of metallic centers and organic ligands that form porous crystalline materials with tunable structures. The vast numbers of available metallic and organic units produce a wide variety of MOFs that can be tailored for applications such as photocatalysis, energy conversion, drug delivery, and microelectronics. Aron Walsh, a professor from the Department of Materials at Imperial College London, commended these findings: "Density functional theory has become a powerful tool for materials modeling. Although we generally understand its performance limitations for organic and inorganic crystals, metal–organic frameworks are less well-understood, which makes this an important study."



Depiction of metal–organic frameworks (MOFs) bandgap data at four levels of theory is represented as raincloud plots. Density functional theory-computed bandgaps were used to generate the Quantum (QMOF) Database using a combination of the Perdew–Burke–Ernzerhof or PBE (GGA, generalized gradient approximation), HLE17 (meta-GGA), HSE06* (10% Hartree–Fock exchange), and HSE06 (25% Hartree–Fock exchange) levels of theory. PBE tends to underpredict MOF bandgaps, whereas HLE17, HSE06, and HSE06* increase the predicted bandgap values.

The QMOF Database accelerates the predictions of MOF bandgaps based on four different density functional approximations. The researchers have combined data from the relatively inexpensive Perdew–Burke–Ernzerhof (PBE) electronic structure properties with smaller amounts of higher accuracy data from one of three functionals (HLE17, HSE06, and HSE06*) using different machine learning models called multi-task and multi-fidelity neural network models to efficiently predict high-quality data for over 20,000 MOFs.

Walsh says, "They show that using a hybrid exchange–correlation functional results in a blueshift in the distribution of MOF bandgaps and that the trends can be reproduced using various statistical models." This breakthrough serves to pave the way for expedited materials discovery and design.

Regarding future work, Rosen says, "Could we predict the thermodynamic properties and chemical stability of metal–organic frameworks? Could we understand and predict their mechanical properties from machine learning? These are all possibilities we can derive from quantum mechanical calculations."

The QMOF Database is available on the widely used Materials Project website: https://materialsproject.org/mofs.

Senam Tamakloe