Energy-Based Descriptors to Rapidly Predict Hydrogen Storage in MOFs

Benjamin J. Bucior[1], N. Scott Bobbitt[1], Timur Islamoglu[2], Subhadip Goswami[2], Arun Gopalan[1], Taner Yildirim[3], Omar K. Farha[2], Neda Bagheri[1], Randall Q. Snurr[1, *]

1. Chemical and Biological Engineering, Northwestern University, Evanston, IL
2. Dept. of Chemistry, Northwestern University, Evanston, IL
3. NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, MD

High-throughput molecular simulations are a powerful tool for MOF design, particularly for adsorption applications like gas storage and separations. These methods can identify promising MOF candidates out of thousands of possibilities, and there are opportunities to better utilize the vast amounts of generated data. We applied machine learning methods to accelerate MOF screening and to develop better structure-property relationships. Our approach uses least absolute shrinkage and selection operator (LASSO) regression to extract insights from MOF energy landscapes, which are rapidly calculated using a probe molecule. The model is highly accurate, interpretable, robust, and three orders of magnitude faster than detailed molecular simulations. To demonstrate the usefulness of this model, we screened 55,000 MOFs and identified a candidate MOF with high hydrogen storage capacity, which we also confirmed experimentally. The approach can be easily applied to other classes of porous materials and adsorbate molecules such as methane.