Supporting Information

Transferable and extensible machine learning derived atomic charges for modeling hybrid nanoporous materials

Vadim V. Korolev,[†,‡,*] Artem Mitrofanov,[†,‡] Ekaterina I. Marchenko,[⊥,§] Nickolay N. Eremin,[§] Valery Tkachenko,[†] Stepan N. Kalmykov[‡]

[†]*Science Data Software, LLC, 14909 Forest Landing Circle, Rockville, Maryland 20850, United States*

[‡]*Department of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia*

[⊥]*Laboratory of New Materials for Solar Energetics, Department of Materials Science, Lomonosov Moscow State University, 119991 Moscow, Russia*

[§]*Department of Geology, Lomonosov Moscow State University, 119991 Moscow, Russia*

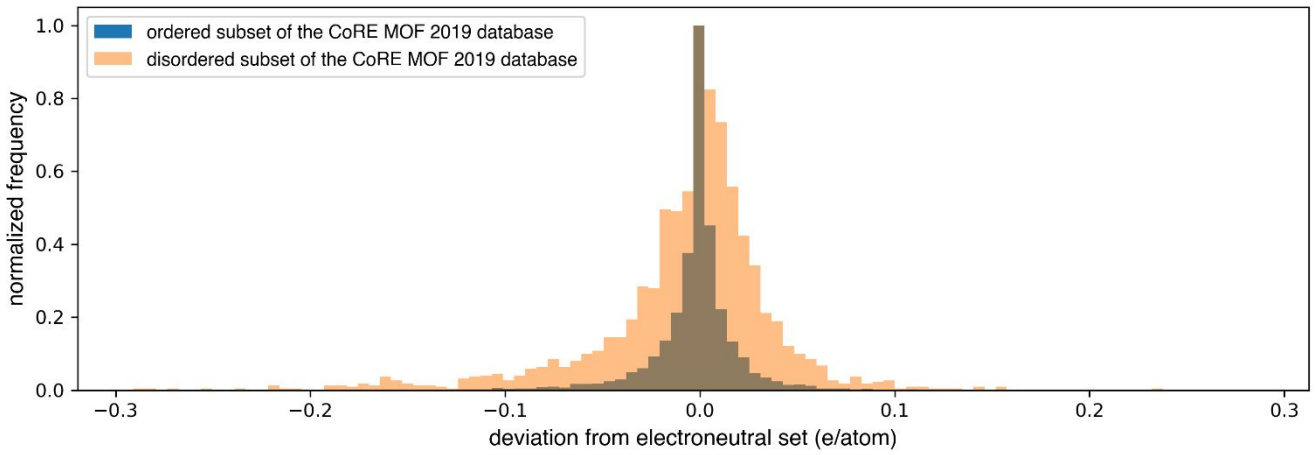

Figure S1. Deviation of predicted ML-derived partial charges from corrected ML charges corresponding to electroneutral unit cell for two subsets of the CoRE MOF 2019 database.

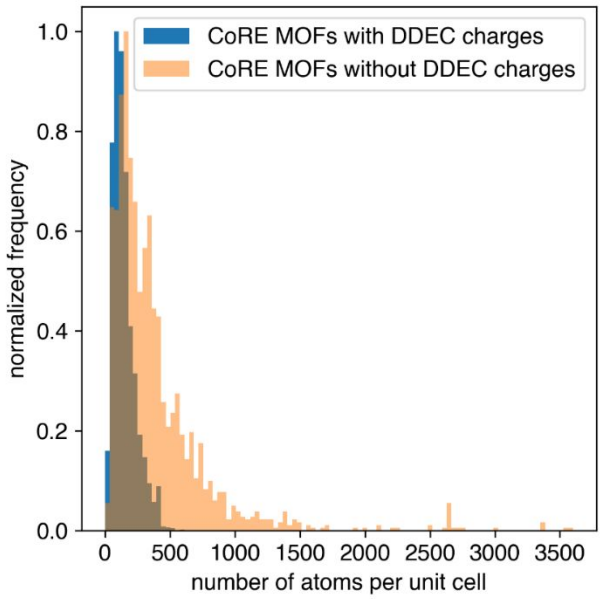

Figure S2. Distribution of structures from the CoRE MOF 2014 database by the number of atoms in the unit cell.

Table S1. Elemental properties.

Atomic number
Atomic weight
Column on periodic table
Covalent radius
DFT bandgap energy of T=0 K ground state
DFT magnetic moment of T=0 K ground state
DFT volume per atom of T=0 K ground state
Melting temperature
Mendeleev Number (position on the periodic table, counting column-wise from H)
Number of filled *d* valence orbitals
Number of filled *f* valence orbitals
Number of filled *p* valence orbitals
Number of filled *s* valence orbitals
Number of unfilled *d* valence orbitals
Number of unfilled *f* valence orbitals
Number of unfilled *p* valence orbitals
Number of unfilled *s* valence orbitals
Number of unfilled valence orbitals
Number of valence electrons
Pauling electronegativity
Row on periodic table
Thermal conductivity
Velocity of sound

Table S2. Structural descriptors.

All the following descriptors were extracted with matminer[1] package.

CoordinationNumber[2]	Number of first nearest neighbors of a site
VoronoiFingerprint[3]	Voronoi indices, i-fold symmetries and statistics of Voronoi facet areas, sub-polyhedron volumes and distances derived by Voronoi tessellation analysis.
CrystalSiteFingerprint[2]	Coordination number percentage and local structure order parameters computed from the neighbor environment of a site; Voronoi decomposition-based neighbor finding.
OPSiteFingerprint[2]	Local structure order parameters computed from the neighbor environment of a site; distance-based neighbor finding.
AGNIFingerprint[4,5]	Fingerprints based on integrating the distances product of the radial distribution function with a Gaussian window function
GaussianSymmFunc[6,7]	Gaussian radial and angular symmetry functions originally proposed for fitting machine learning potentials.

References

(1) Ward, L.; Dunn, A.; Faghaninia, A.; Zimmermann, N. E. R.; Bajaj, S.; Wang, Q.; Montoya, J.; Chen, J.; Bystrom, K.; Dylla, M.; others. Matminer: An Open Source Toolkit for Materials Data Mining. *Comput. Mater. Sci.* **2018**, *152*, 60–69.

(2) Zimmermann, N. E. R.; Horton, M. K.; Jain, A.; Haranczyk, M. Assessing Local Structure Motifs Using Order Parameters for Motif Recognition, Interstitial Identification, and Diffusion Path Characterization. *Front. Mater.* **2017**, *4*, 1–13.

(3) Boots, B.; Okabe, A.; Sugihara, K. *Spatial Tessellations*; John Wiley & Sons New York, NY, 1999; Vol. 1.

(4) Botu, V.; Ramprasad, R. Learning Scheme to Predict Atomic Forces and Accelerate Materials Simulations. *Phys. Rev. B* **2015**, *92* (9), 94306.

(5) Botu, V.; Ramprasad, R. Adaptive Machine Learning Framework to Accelerate Ab Initio Molecular Dynamics. *Int. J. Quantum Chem.* **2015**, *115* (16), 1074–1083.

(6) Khorshidi, A.; Peterson, A. A. Amp: A Modular Approach to Machine Learning in Atomistic Simulations. *Comput. Phys. Commun.* **2016**, *207*, 310–324.

(7) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134* (7), 74106.