Computational Screening of MOFs for Carbon Capture

Samir Budhathoki[1,2], Christopher E. Wilmer[3], Jan Steckel[1]

[1]National Energy Technology Laboratory, 626 Cochran Mill Road, Pittsburgh, PA 15236, USA

[2]NETL Support Contractor, 626 Cochran Mill Road, Pittsburgh, PA 15236, USA,

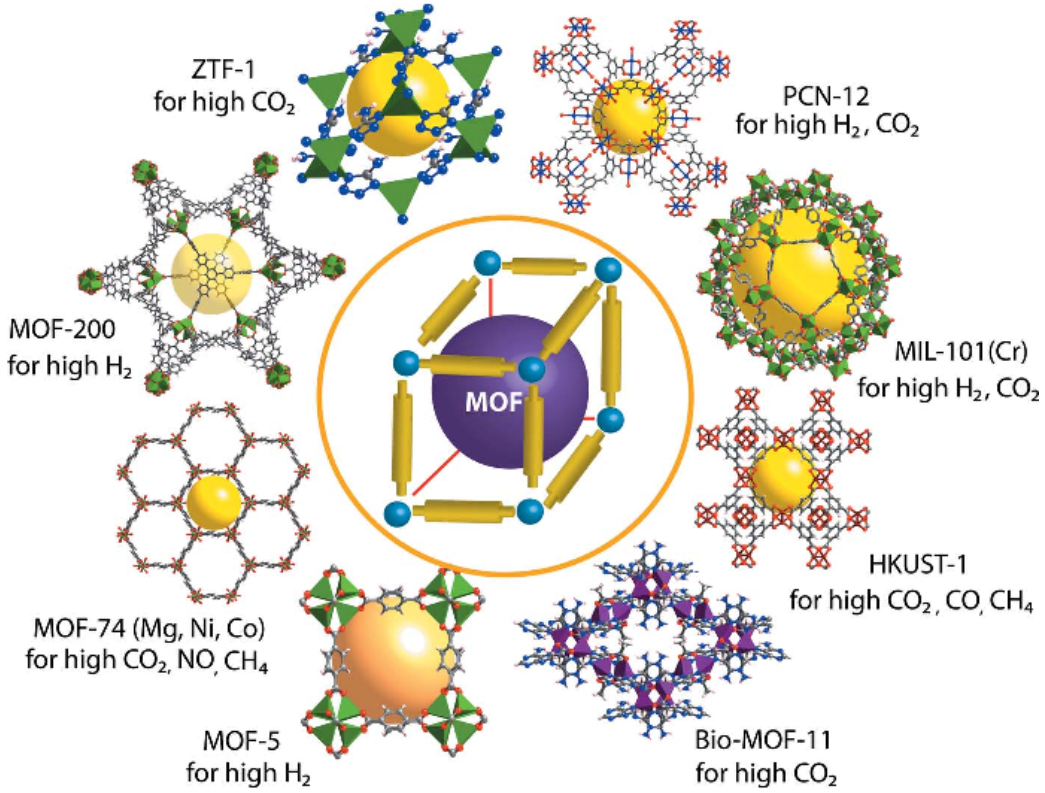
[3]University of Pittsburgh, Pittsburgh, PA 15260, USA

Disclaimer

This project was funded by the United States Department of Energy, National Energy Technology Laboratory, in part, through a site support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Objective

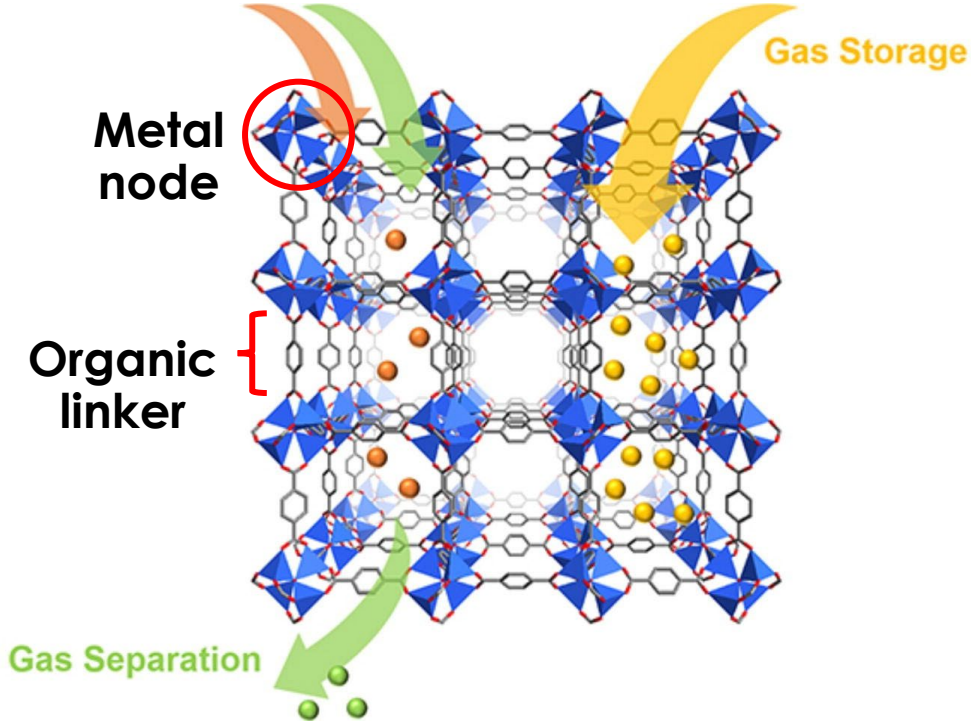
- Develop computational methods for accurately screening CO_2 sorption in MOF materials.
- The desired products of this task are identification of MOFs with high CO_2 loading at low pressure



Zhao et al. *Energy Storage Materials* **2016**, *2*, 35-62.

Metal Organic Frameworks (MOFs)

- Crystalline, porous materials
- Good for gas storage/separation applications
- Large, diverse class of molecules
  - >100,000 synthesized
  - >500,000 predicted



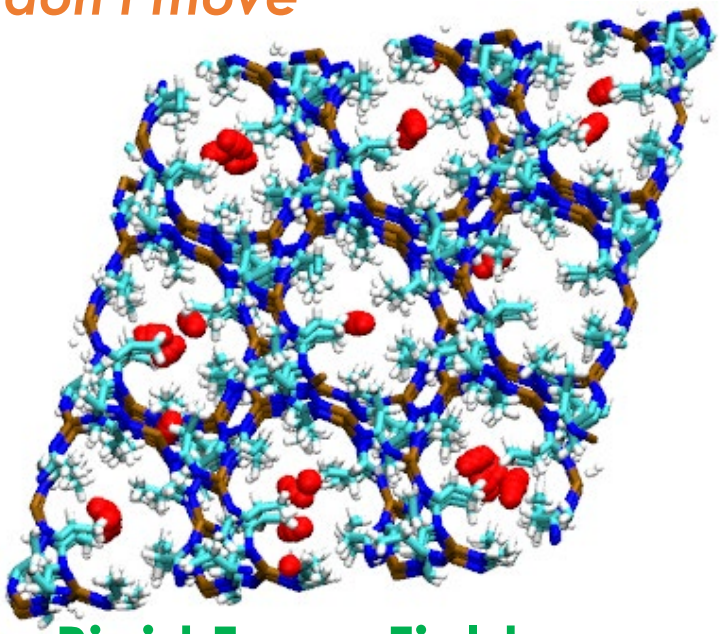
Li et al. *Materials Today* **2018**, *21* (2), 108–121.

MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs as rigid - MOF atoms don't move

MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs as rigid - MOF atoms don't move

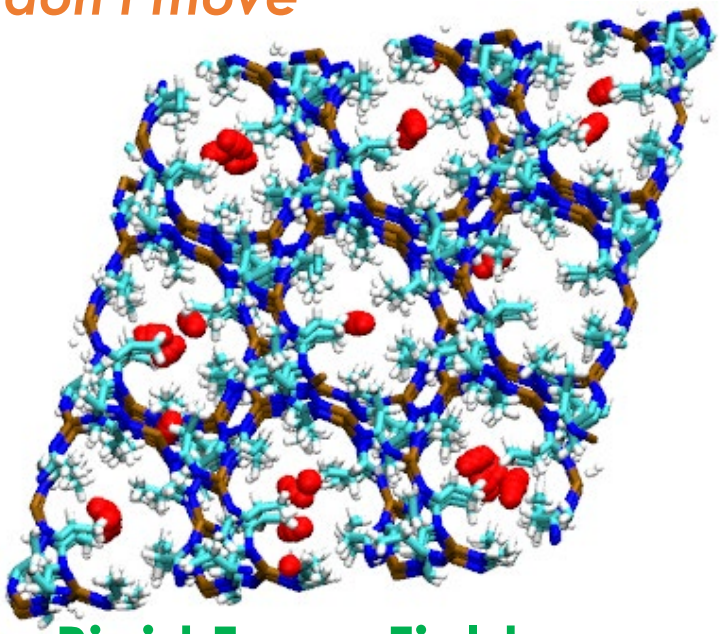


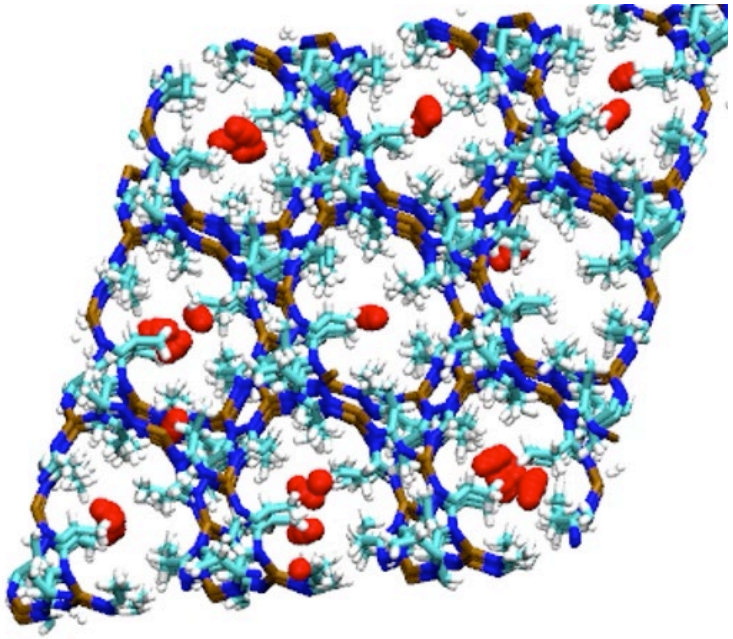
Rigid Force Field

- Electrostatic + dispersion
- low computational cost
- Good for rigid materials but most of the MOFs are flexible
- Easy to obtain

MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs as rigid - MOF atoms don't move



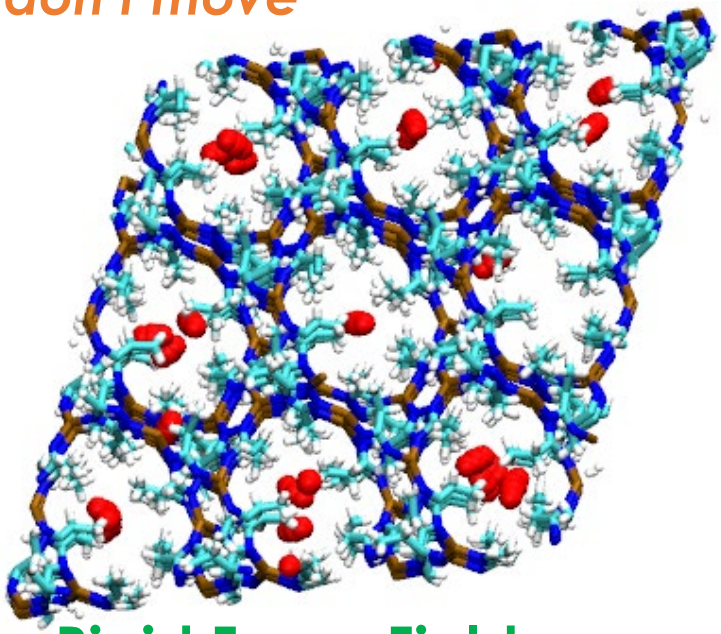


Rigid Force Field

- Electrostatic + dispersion
- low computational cost
- Good for rigid materials but most of the MOFs are flexible
- Easy to obtain

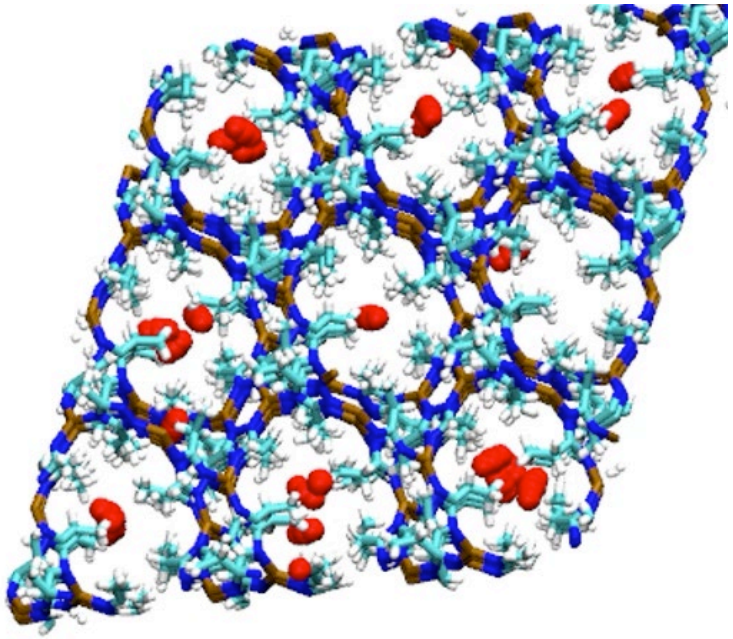
MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs as rigid - MOF atoms don't move



Rigid Force Field

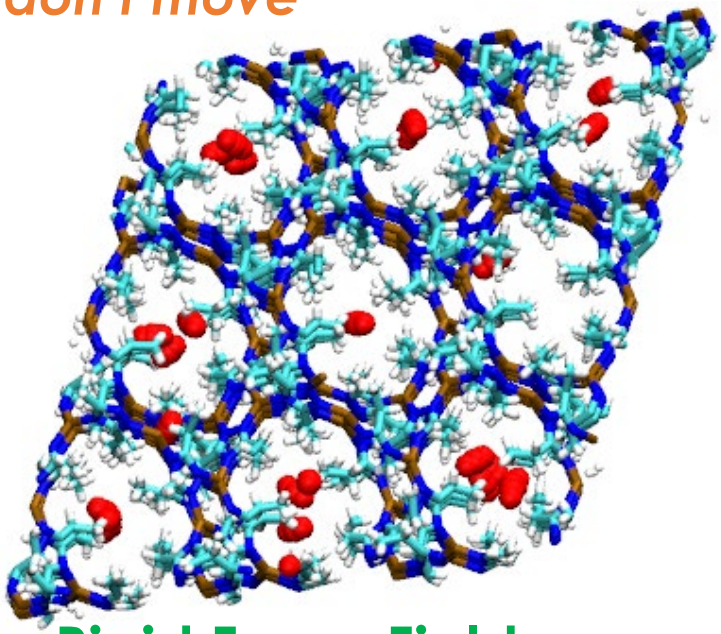
- Electrostatic + dispersion
- low computational cost
- Good for rigid materials but most of the MOFs are flexible
- Easy to obtain



Flexible Force Field

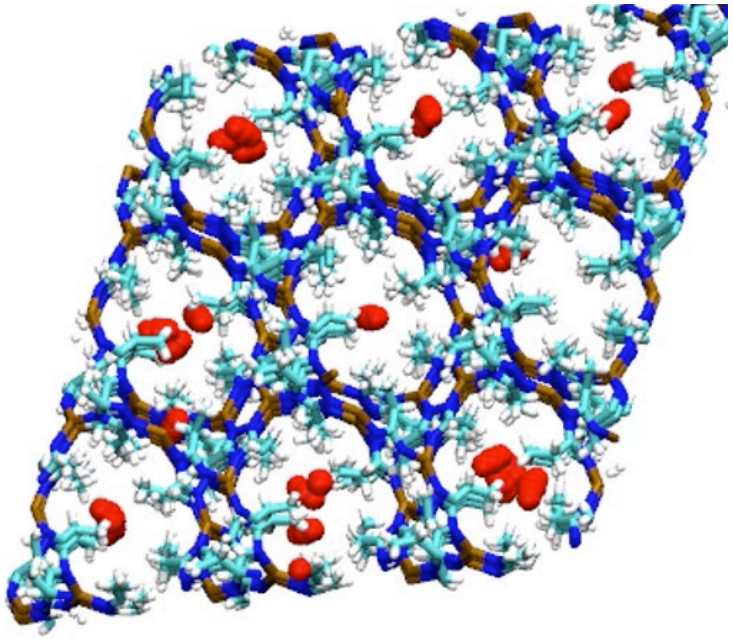
MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs as rigid - MOF atoms don't move



Rigid Force Field

- Electrostatic + dispersion
- low computational cost
- Good for rigid materials but most of the MOFs are flexible
- Easy to obtain

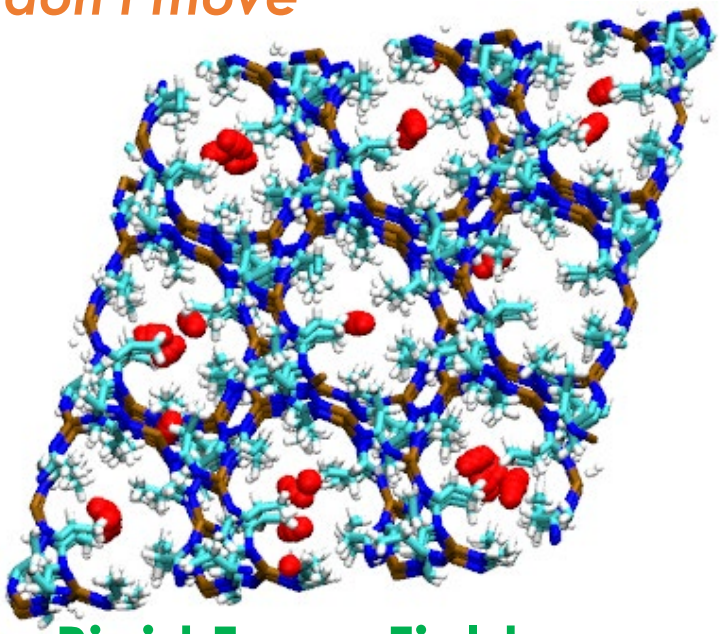


Flexible Force Field

- Electrostatic + dispersion + bond + angle + torsion

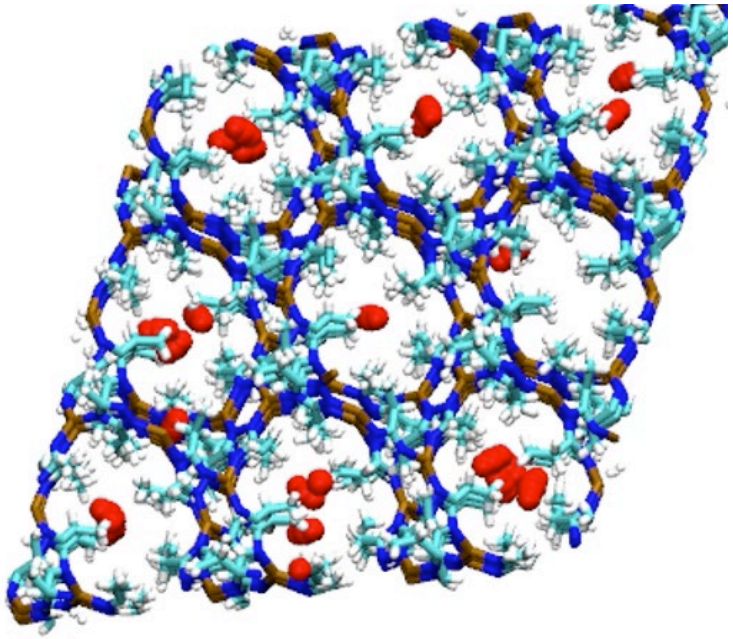
MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs as rigid - MOF atoms don't move



Rigid Force Field

- Electrostatic + dispersion
- low computational cost
- Good for rigid materials but most of the MOFs are flexible
- Easy to obtain

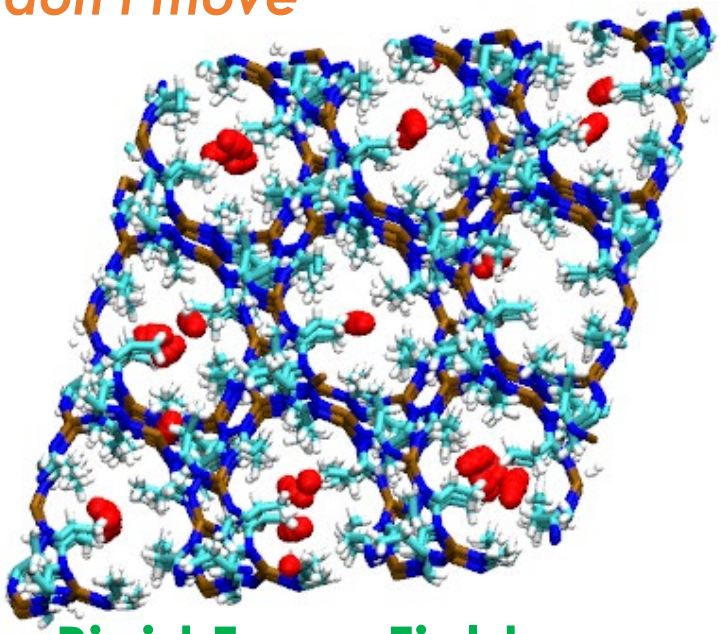


Flexible Force Field

- Electrostatic + dispersion + bond + angle + torsion
- High computational cost

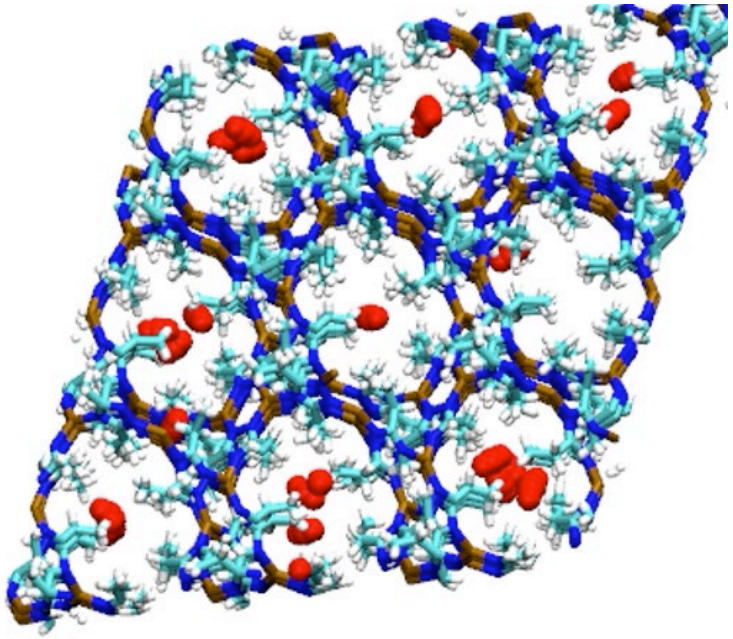
MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs as rigid - MOF atoms don't move



Rigid Force Field

- Electrostatic + dispersion
- low computational cost
- Good for rigid materials but most of the MOFs are flexible
- Easy to obtain

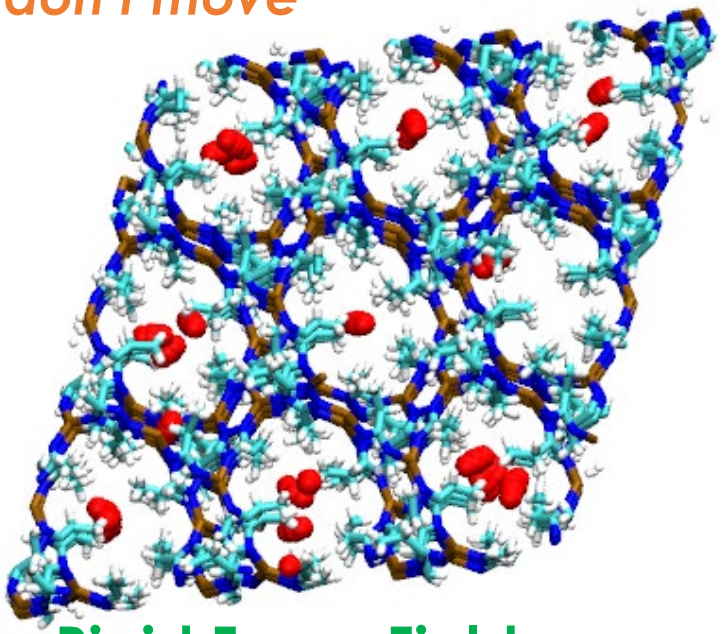


Flexible Force Field

- Electrostatic + dispersion + bond + angle + torsion
- High computational cost
- Accounts for MOF flexibility

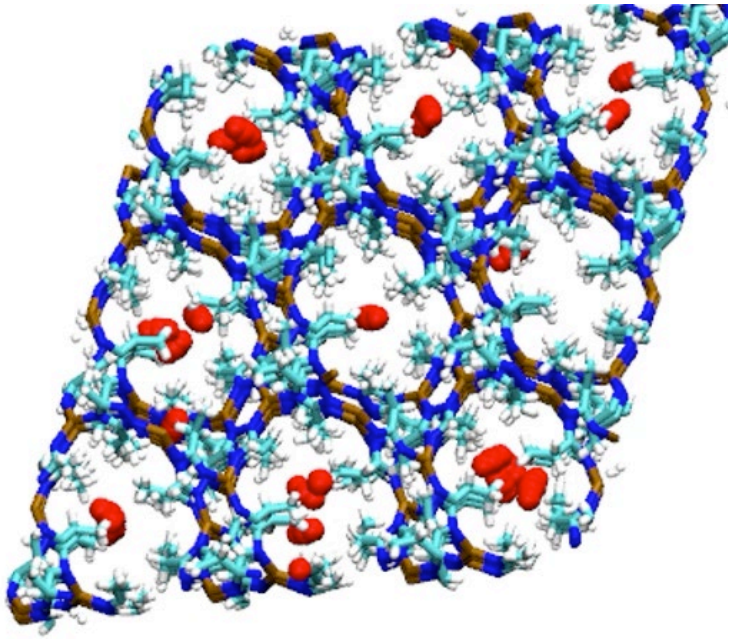
MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs as rigid - MOF atoms don't move



Rigid Force Field

- Electrostatic + dispersion
- low computational cost
- Good for rigid materials but most of the MOFs are flexible
- Easy to obtain

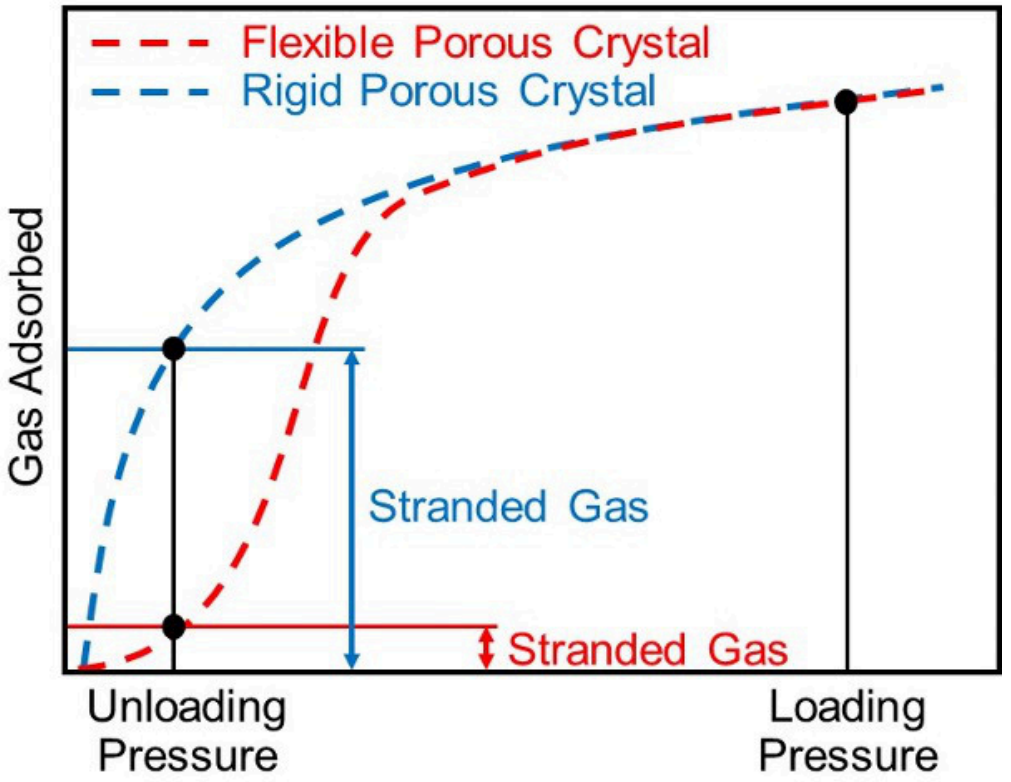


Flexible Force Field

- Electrostatic + dispersion + bond + angle + torsion
- High computational cost
- Accounts for MOF flexibility
- Very hard to obtain for a large set of materials

We take account of the MOF flexibility

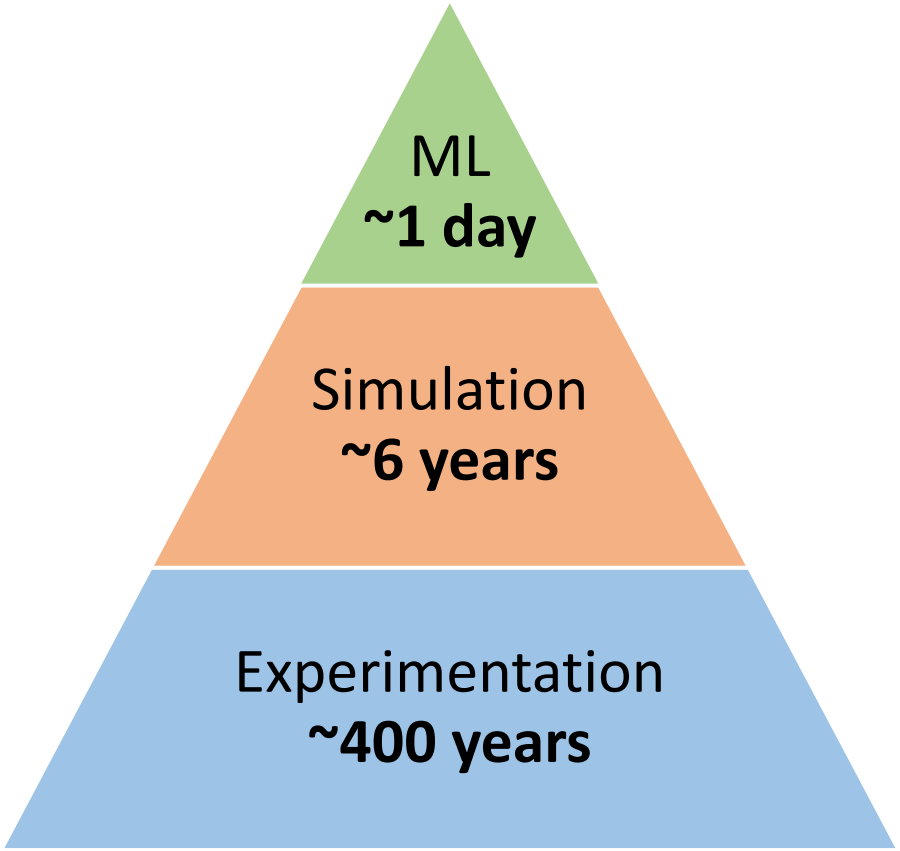
Recent studies have shown that both subtle flexible modes and large framework volume changes can influence MOF performance regarding adsorption-based applications



Mason et. al., *Nature* **2015,** 527, 357-361
Simon et. al., *Proc. Natl. Acad. Sci. U. S. A.* **2017,** 114, E287-E296
Witman et. al., *J. Am. Chem. Soc.* **2017**, 139, 5547-5557
Gladysiak et. al., *ACS Appl. Matter. Interfaces* **2018**, 10, 36144-36156
Heinen et. al., *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, 8, e1363
Jawahery et. al., *J. Chem. Theory Comput* **2019**, 15, 3666-3677

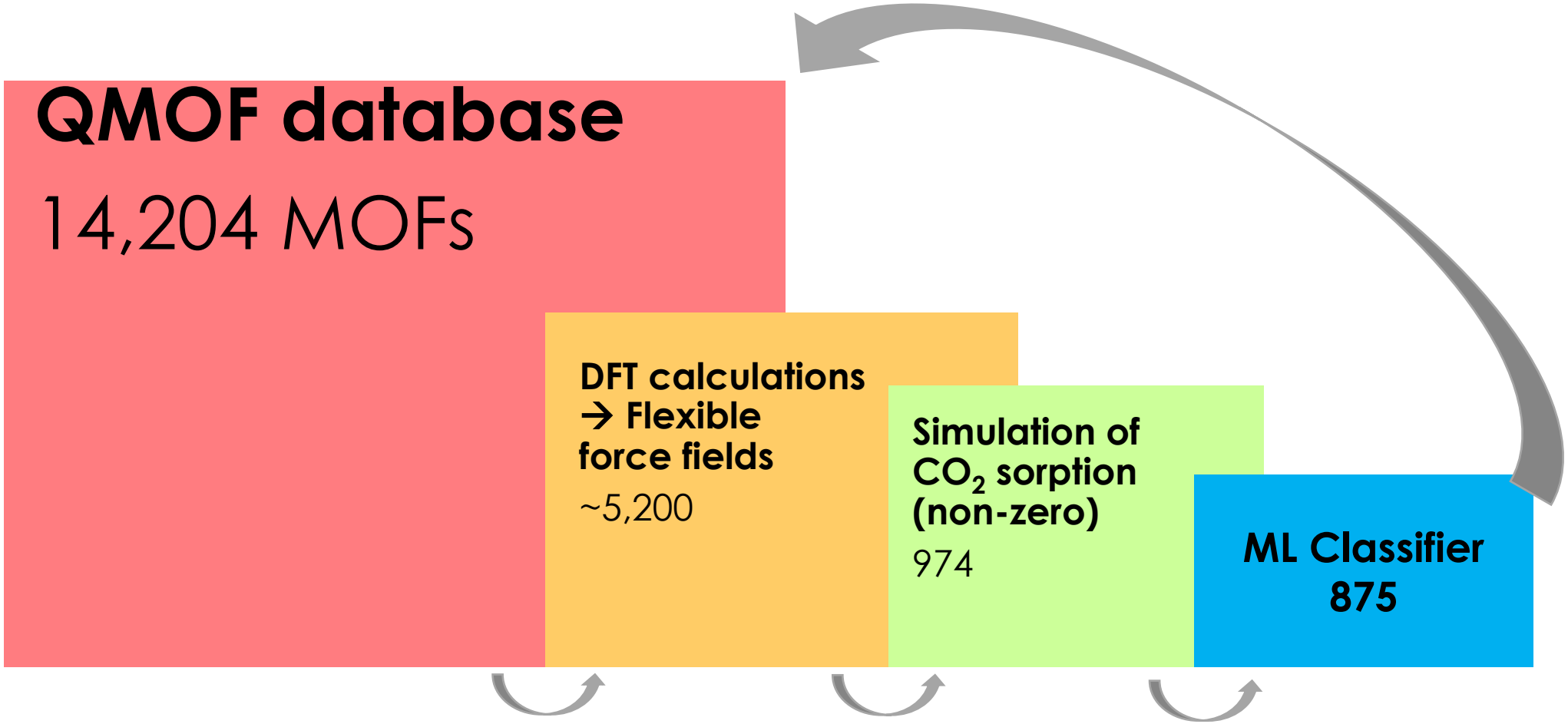
Motivation for Machine Learning

- **ML is faster + easier than traditional methods for evaluating material properties**
- **Computational simulation**
  - long computation times
  - high level of theory
- **Lab experimentation**
  - long experiments
  - unavailable equipment/reagents



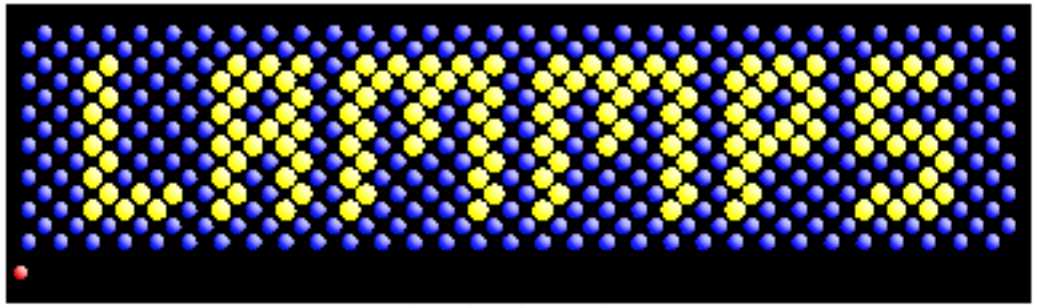
To Screen 500,000 MOFs

Project Approach



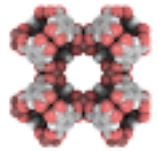
Rosen et. al., *Matter* **2021,** 4, 1-20
Vanduyfhuys et. al., *J Comput Chem* **2018,** 39(16), 999-1011

Software

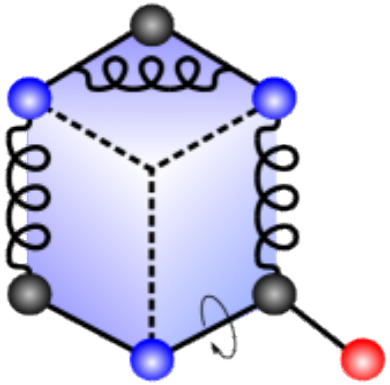


iRASPA/**RASPA2**

Classical molecular simulation code



zeoplusplus

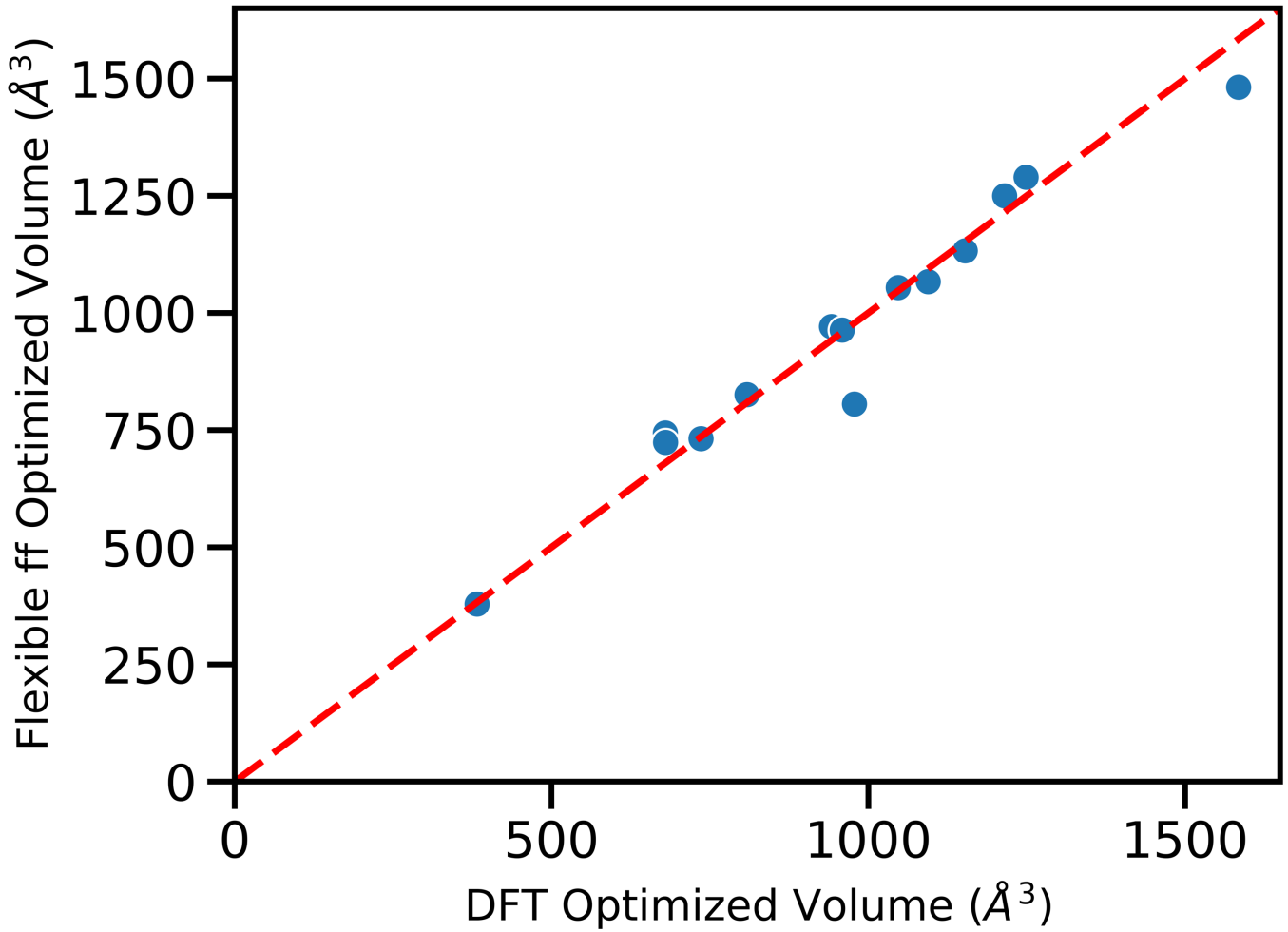


QuickFF

VASP

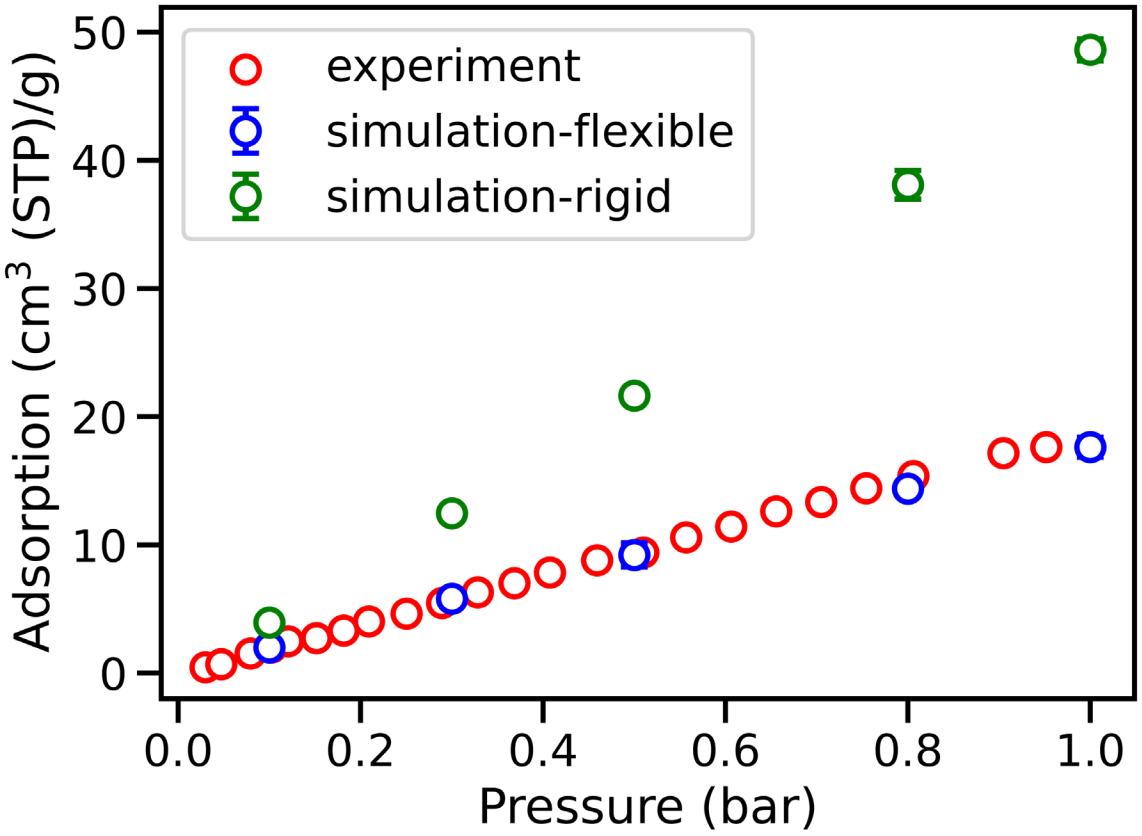
PYCARET

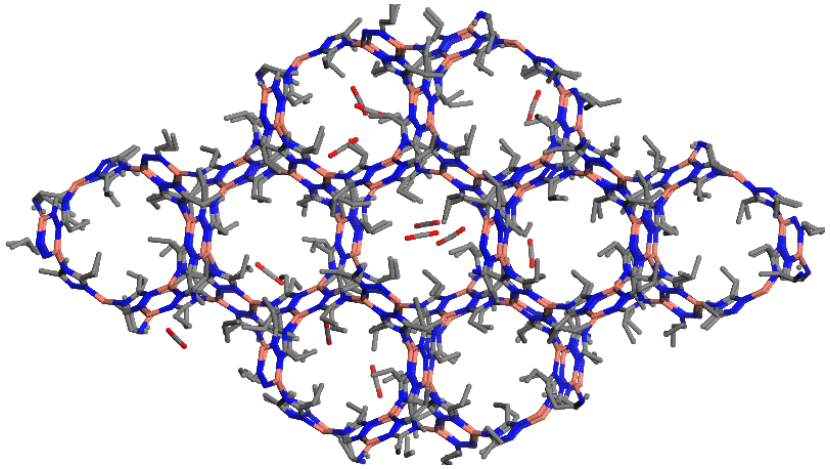
Flexible Force Field Accuracy Validation



Comparison with Experiment

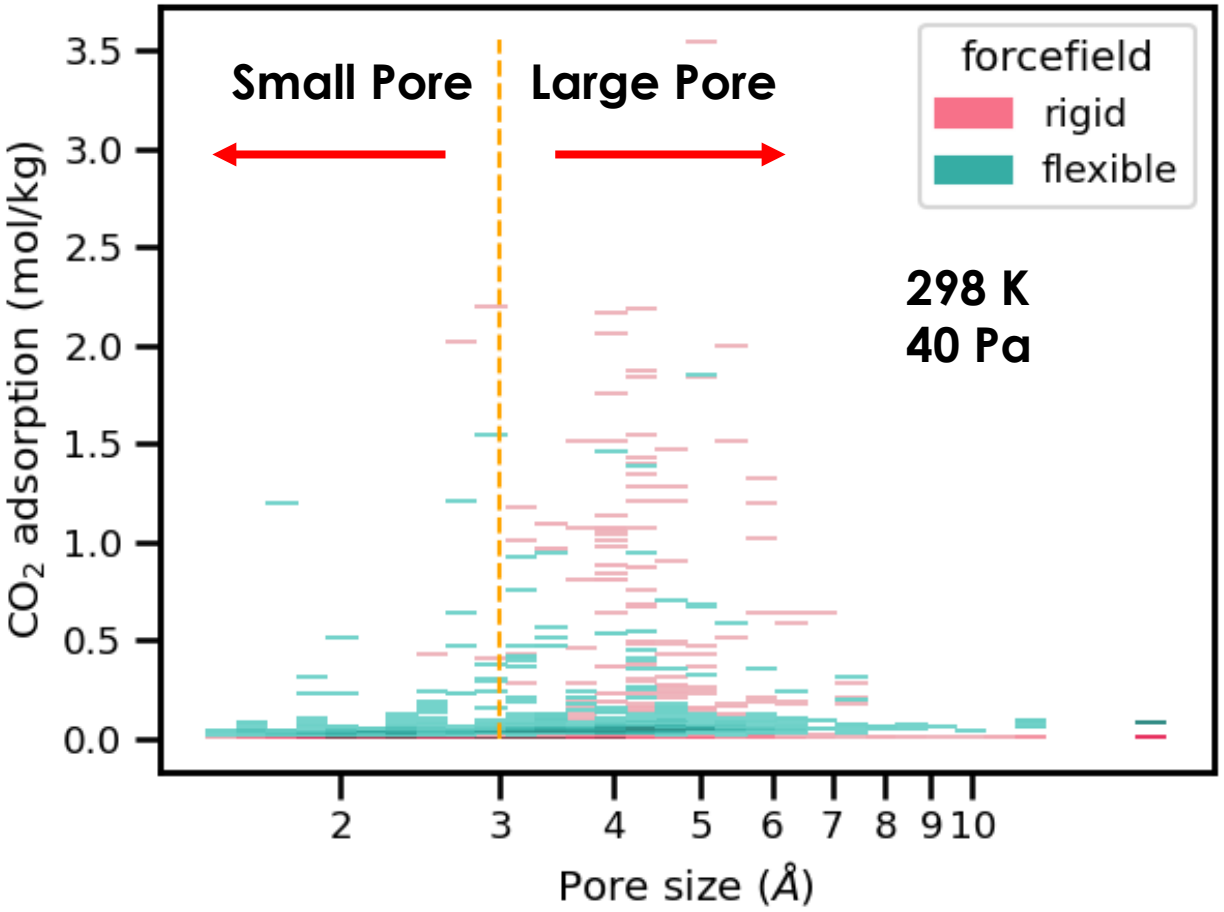
CO_2 adsorption in MAF-2 (BOGXIF) at 298 K





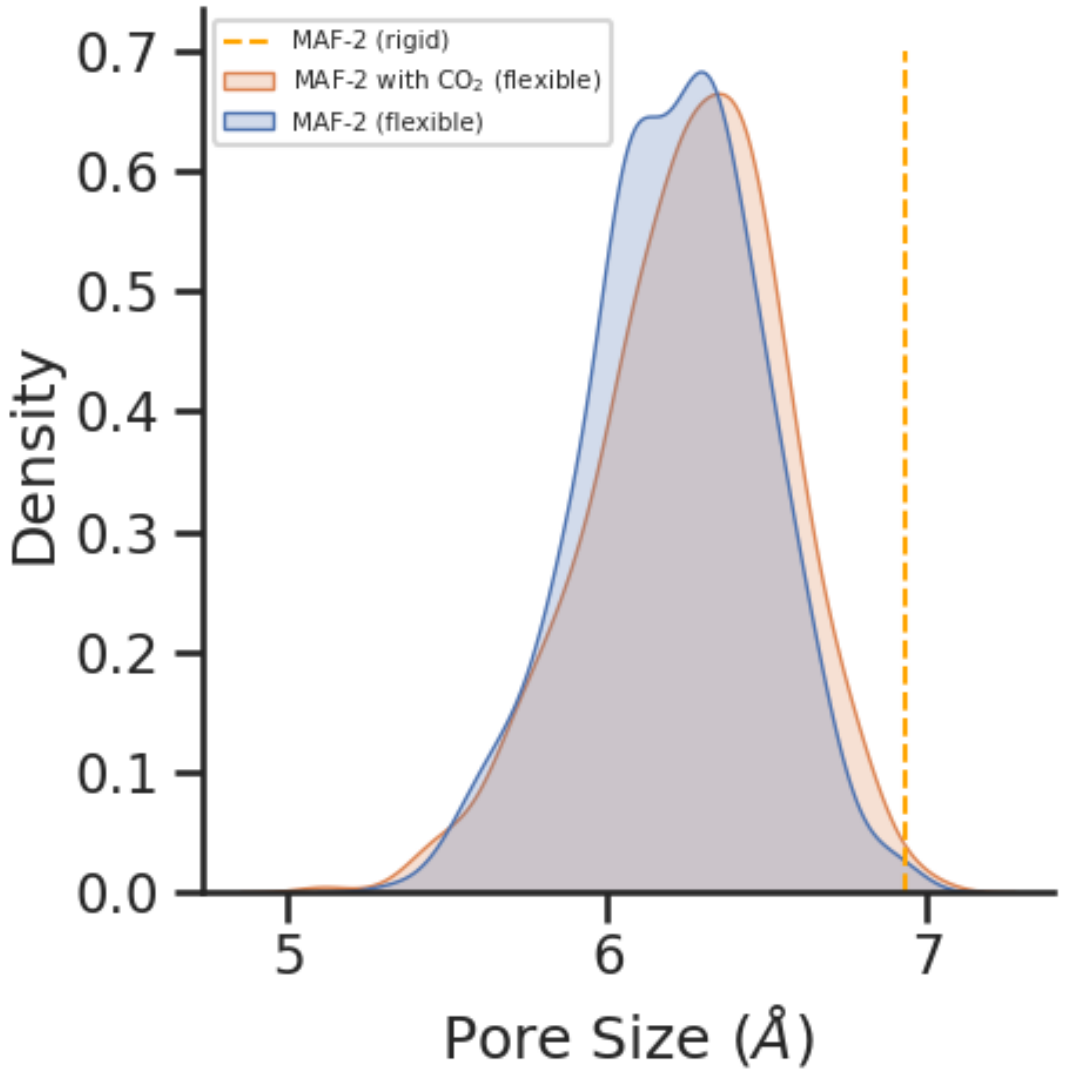
Zhang, J. P.; Chen, X-M.; *J. Am. Chem. Soc.* **2009**, 131, 15, 5516-5521

CO_2 Adsorption at Direct Air Capture Conditions



- Flexible forcefield yield CO_2 adsorption in MOFs with small pore sizes
- Rigid forcefield yield no CO_2 adsorption for MOFs with pore size less than 3 Å in most of the cases
- For some MOFs with pore size greater than 3 Å, rigid forcefield overestimate the CO_2 adsorption

Flexibility Changes Pore Size



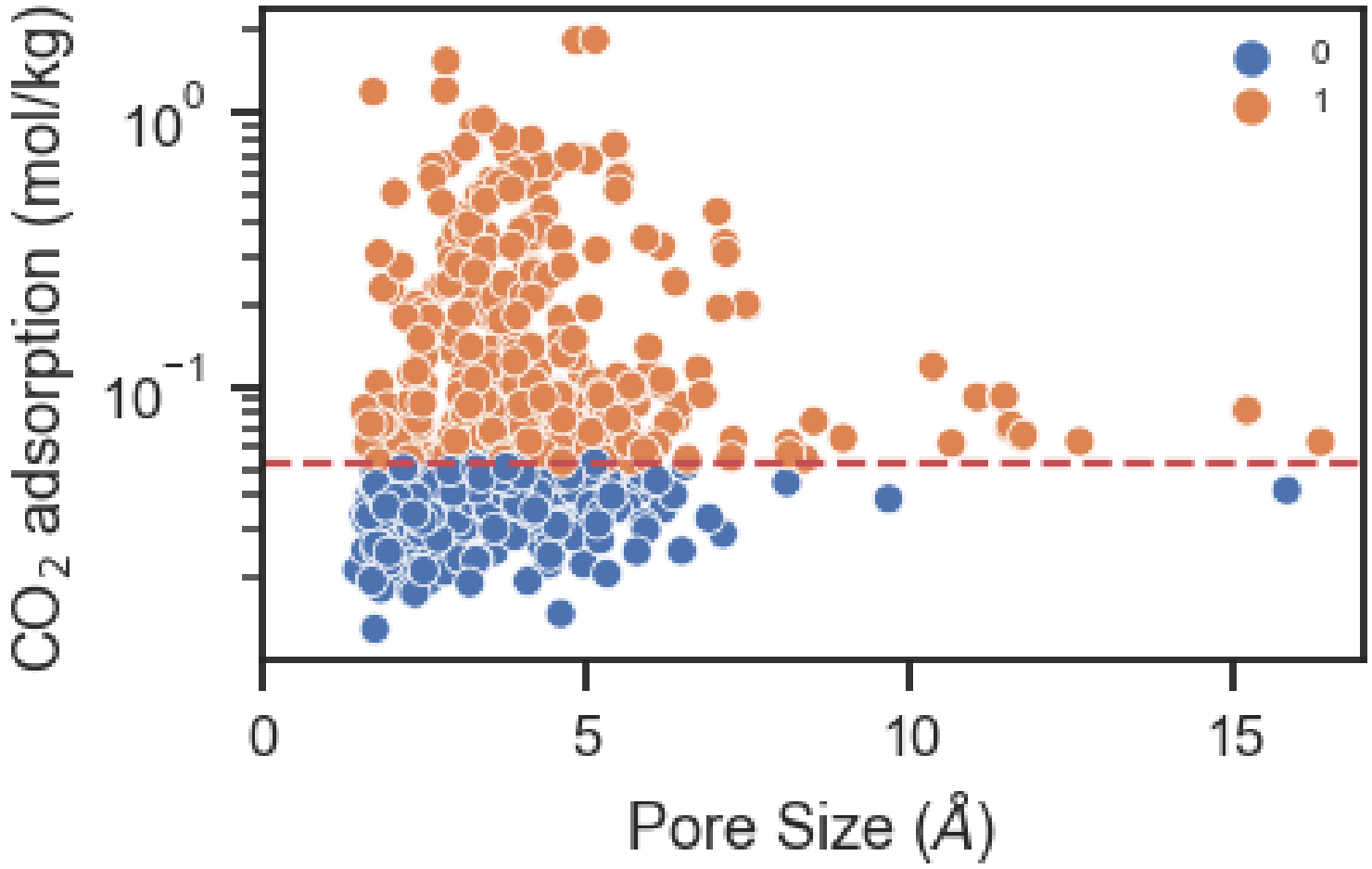
- Presence of CO_2 expands pore
- Distribution of pore size (LCD) is observed for flexible model
- Expansion and contraction of pore play critical roles in CO_2 adsorption

Top 10 CO_2 adsorbing MOFs: Flexible vs Rigid MOF Models

Rank	Flexible	Rigid
1	OF	WA
2	EZ	DI
3	DI	DU
4	DU	DI
5	ME	LA
6	HO	PI
7	DI	EP
8	UJ	PI
9	ME	NE
10	PE	DI

Different Answers!

Machine Learning model for MOF classification

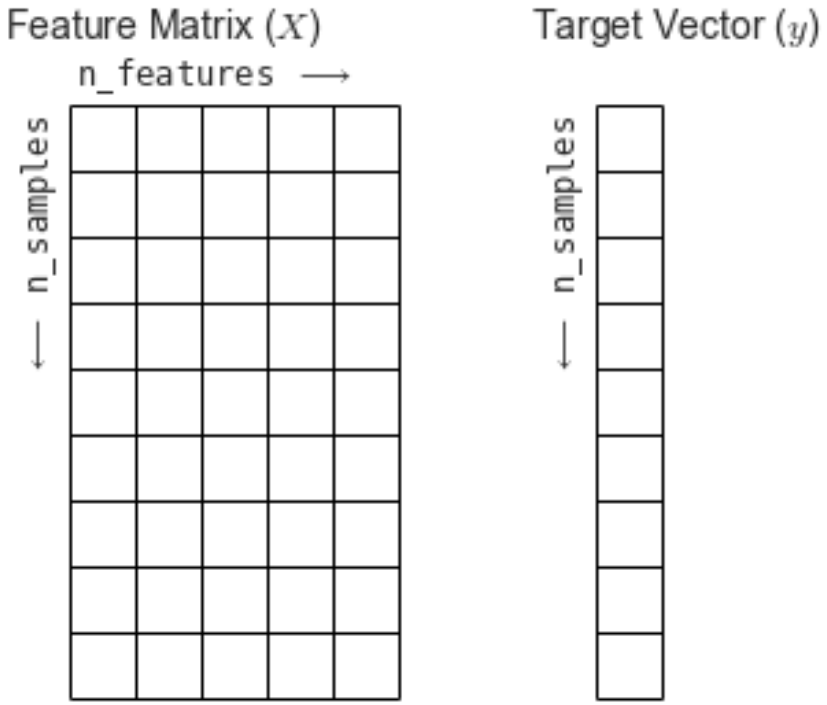


0 – CO_2 ads < 0.052 mol/kg
1 – CO_2 ads >= 0.052 mol/kg

MOF Featurization

Feature: n-dimensional numerical vector that represents each MOF

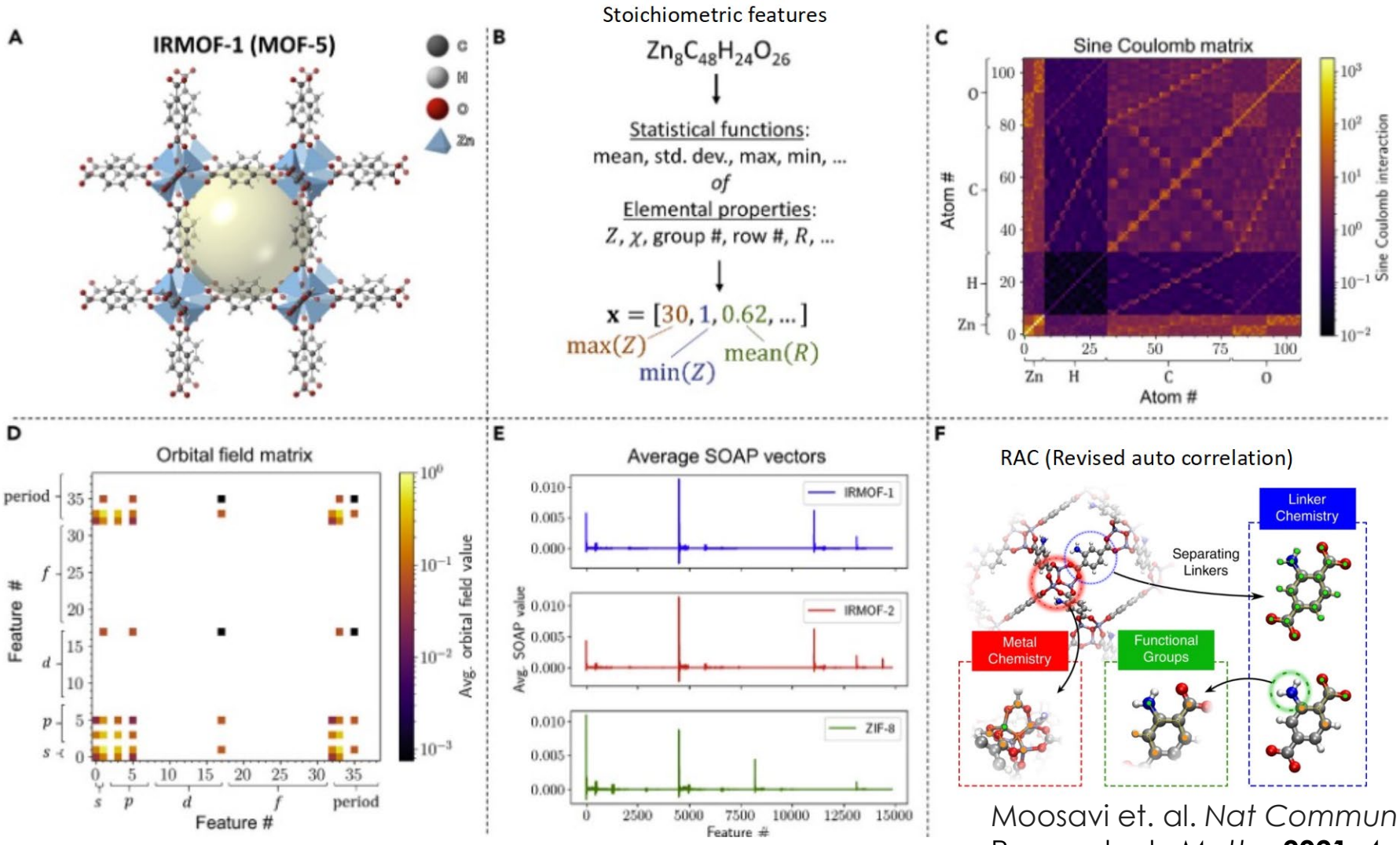
1. **Stoichiometric-45**[1]
   - 45 statistical attributes of elemental properties
2. **Stoichiometric-120**[1]
   - 103 attributes describing elemental fractions
   - 7 statistical attributes of elemental properties
3. **Sine Coulomb Matrix**[1]
   - pairwise electrostatic interactions between nuclei
4. **Orbital Field Matrix**[1]
   - distribution of valence electrons
   - interaction of valence subshells between atoms
5. **Smooth Overlap of Atomic Positions (SOAP)**[1]
   - similarity between a pair of local atomic environments
6. **Revised Autocorrelation (RAC) values + Custom features**
   - molecular revised autocorrelation (RAC) values, surface area, volume, density, pore-limiting diameter (PLD), charge difference, epsilon



jakevdp.github.io

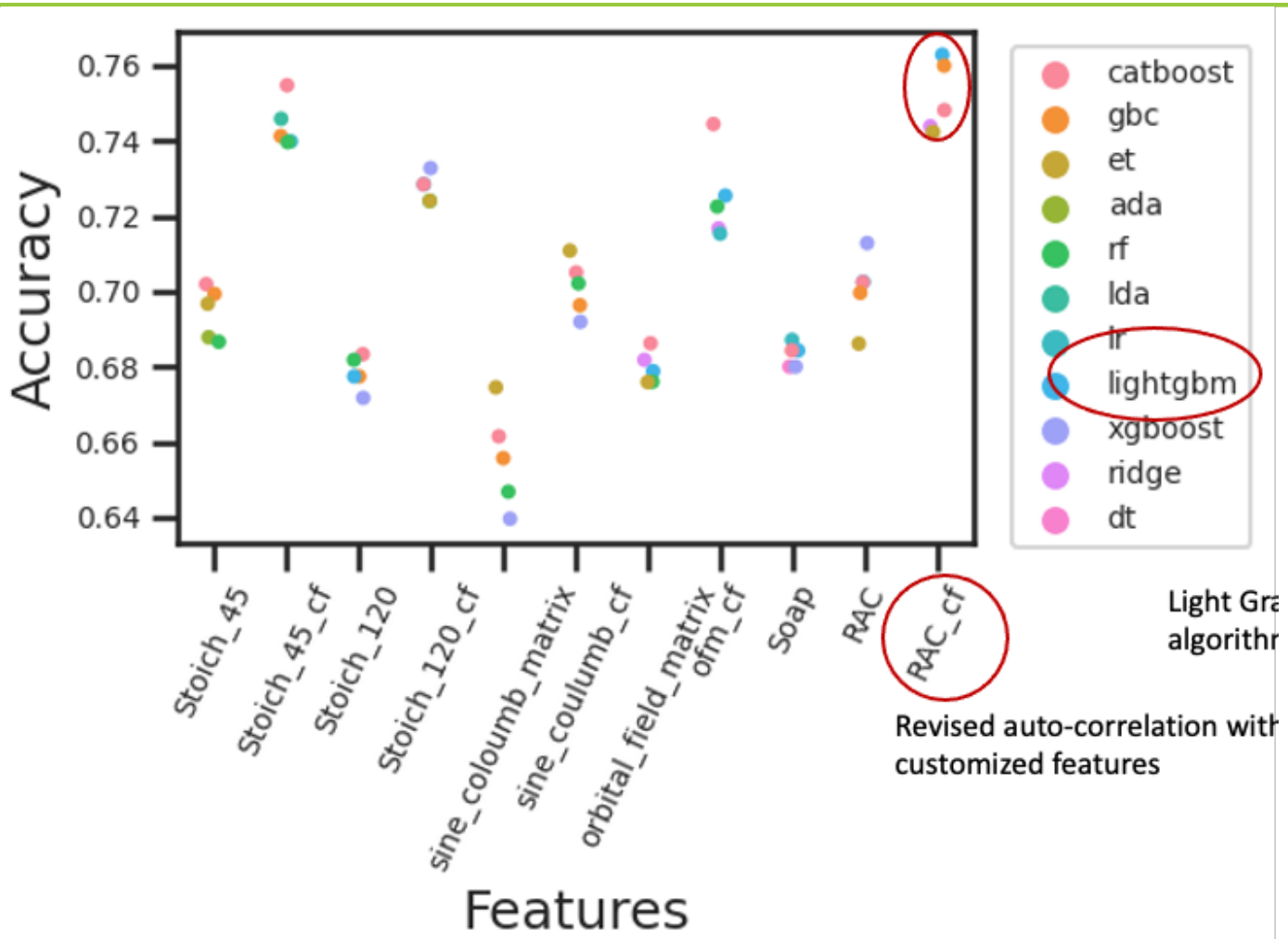
[1] *Matter* **2021**, *4* (5), 1578–1597.

MOF Features



Moosavi et. al. *Nat Commun*, **2020**, 4068
Rosen et. al., *Matter* **2021,** 4, 1-20

Results: ML Classifiers validation results



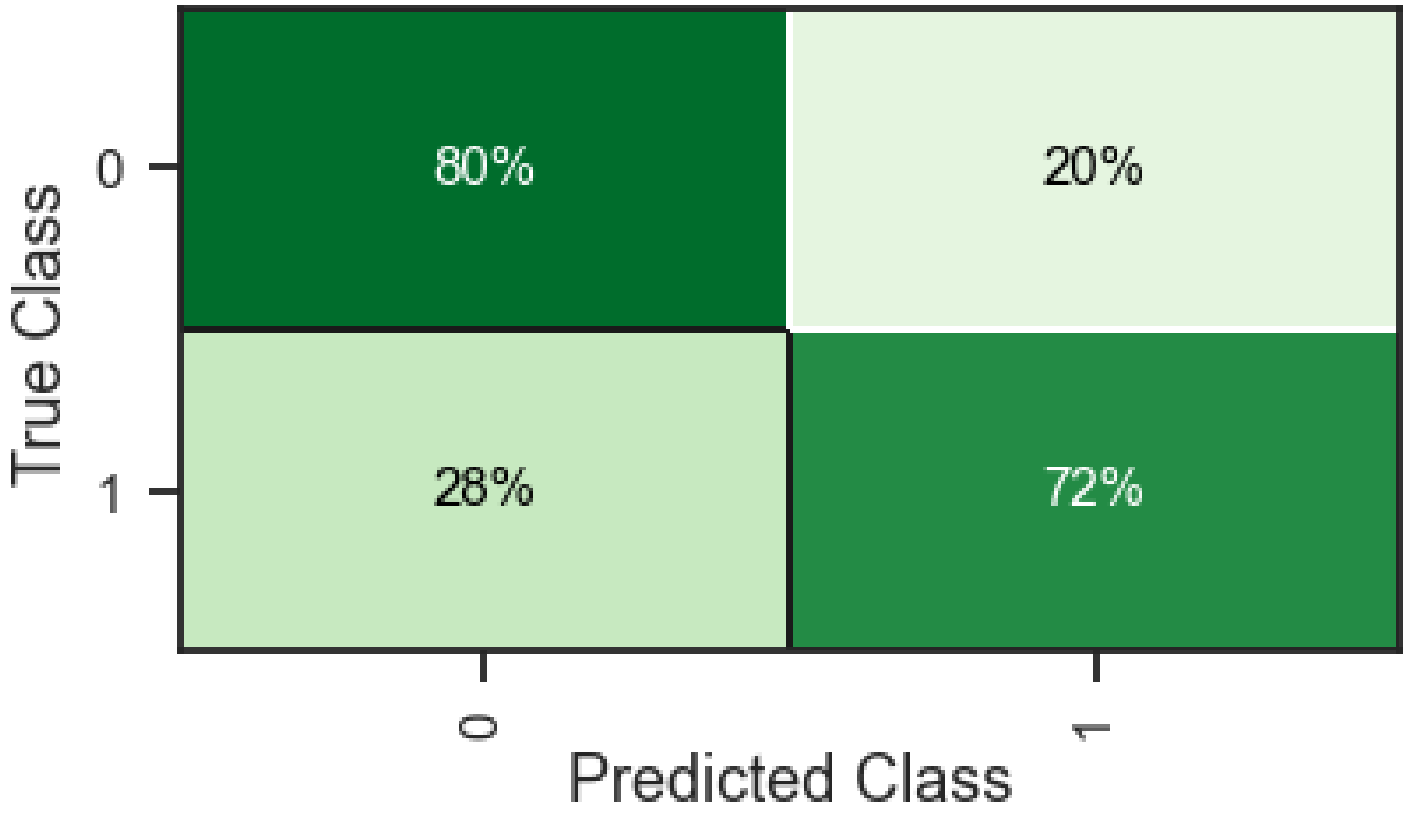
Train size : 80 %
Test size: 20%
Results are averaged from 10 cv folds

- 11 different classification ML models
- All the features were customized (denoted by '_cf' in plot) with geometric features such as surface area, pore size, electrostatics and dispersion term.
- Light Gradient Boosting algorithm performed the best

$$\text{Accuracy} = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + False\ Negative} = 76\%$$

Results: Light Gradient Boosting Classification

Confusion Matrix for Test Set



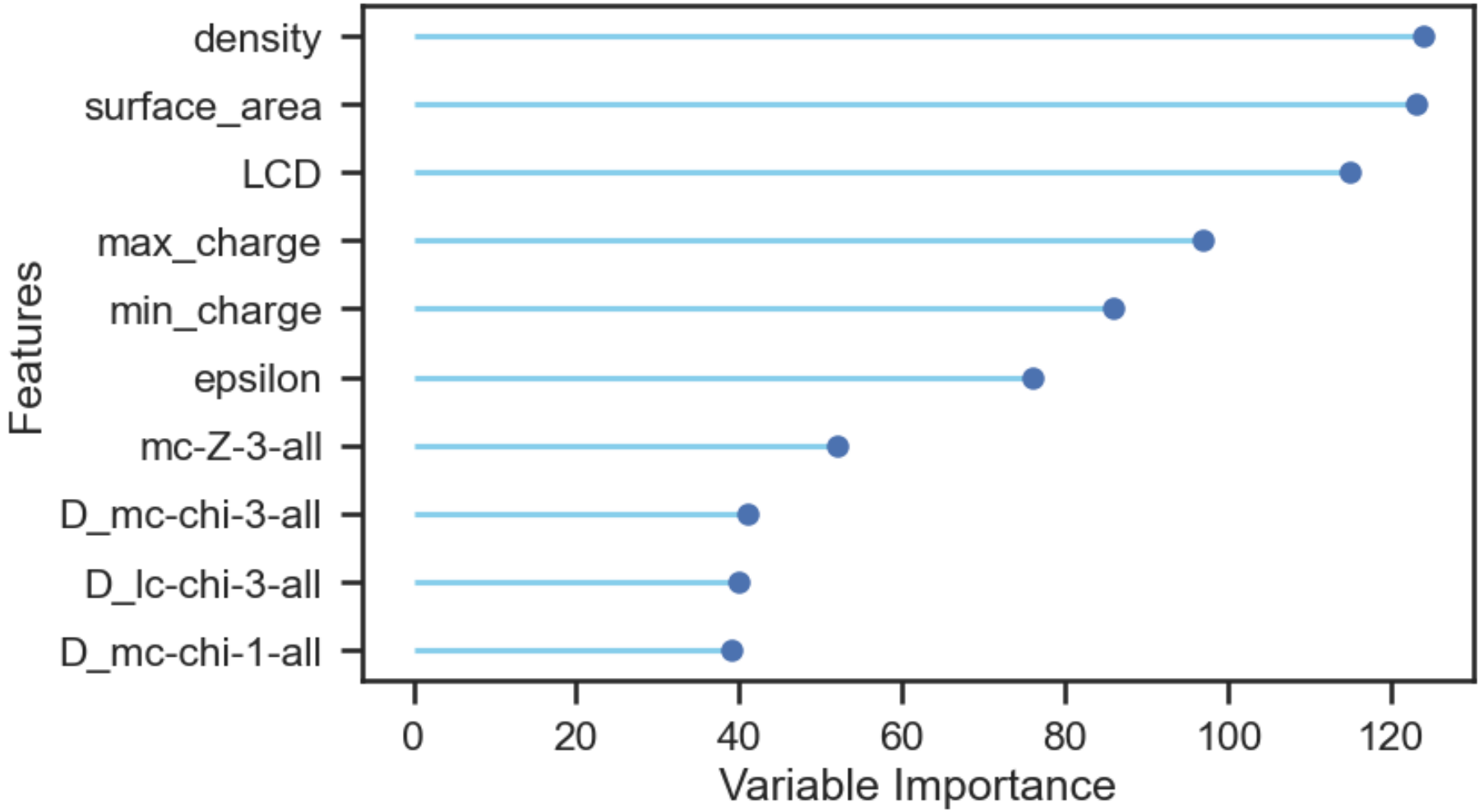
0 : low CO_2 adsorbing MOFs

1 : High CO_2 adsorbing MOFs

$$\text{Accuracy} = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive_{+} True\ Negative + False\ Negative} = 75\%$$

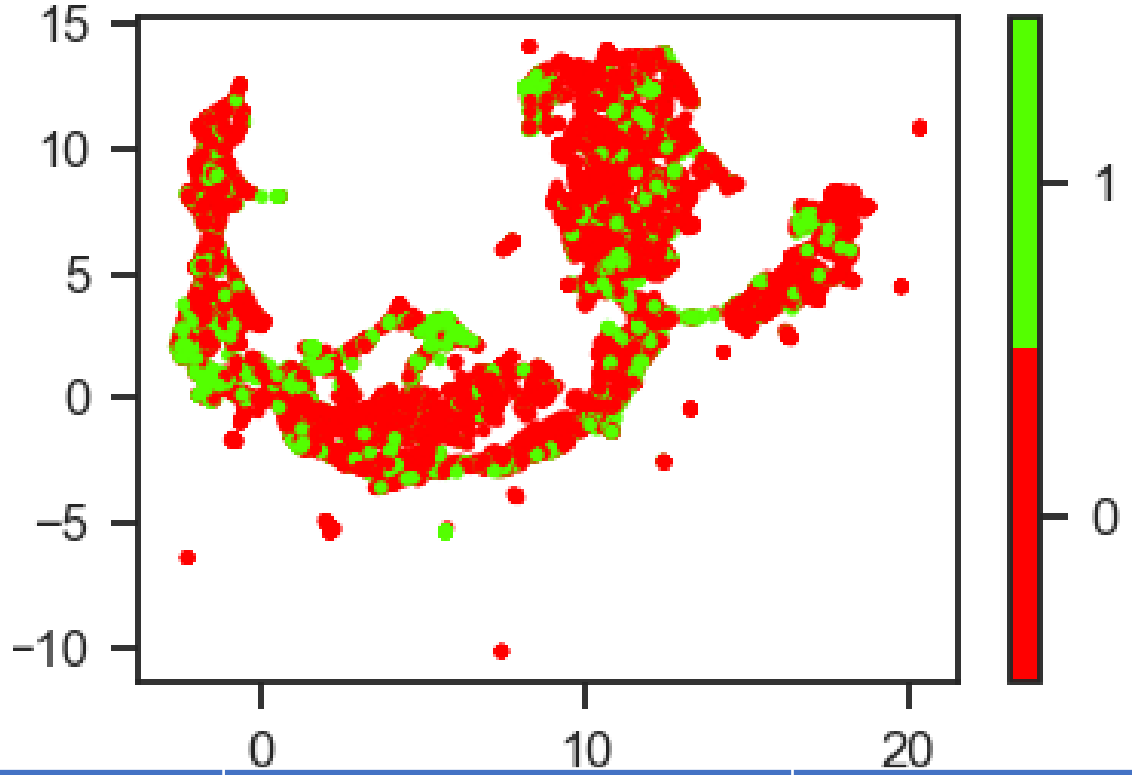
Results: LightGBM Classification

Feature Importance: Top 10 out of 173 Features



- LCD : Pore Size
- Epsilon : Summation of van der Waals interaction
- Mc-Z-3-all: Nuclear charge of metallic center
- D_mc_chi-3-all and D_mc_chi-1-all: Pauling Electronegativities of metallic center
- D_lc-chi-3-all : Pauling electronegativity of linkers

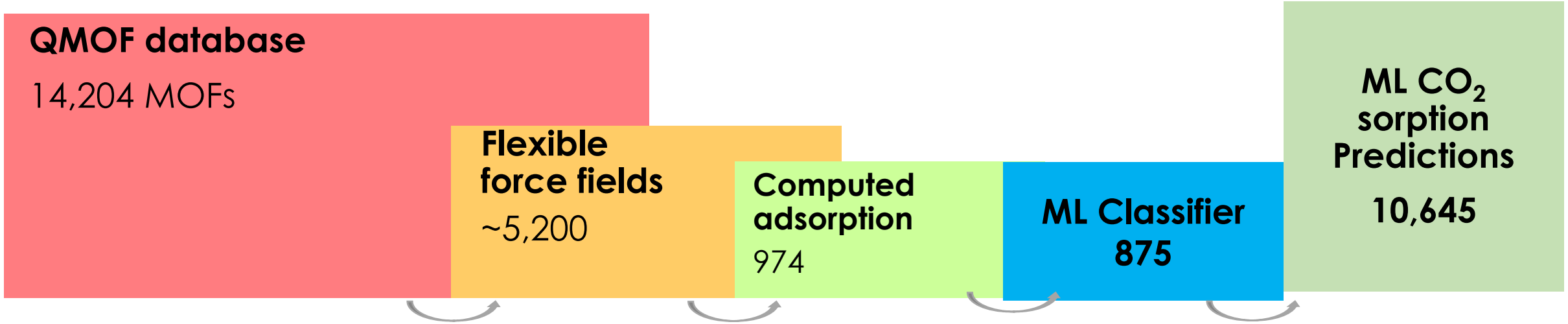
Results: Predictions on rest of QMOF Database



Total MOFs	High adsorbing MOFs (1) (CO_2 adsorption > = 0.05 mol/kg)	Low adsorbing MOFs (0) (CO_2< 0.05 mol/kg)
10,645	2447 (~23 %)	8198 (~77%)

Summary and Conclusions

- Flexible force fields yielded CO_2 sorption for MOFs with pore size < 3 Å
- Ranking of MOFs based on CO_2 sorption was different for flexible model and rigid models.
- Light Gradient Boosting classification Machine Learning (LGBM) yielded about 75% Accuracy
- MOF geometric features (pore size, density, surface area) along with electrostatic and dispersion terms were found to be the important features for ML model
- ML classification model has been used to screen the rest of the 14K QMOF database and can be used to screen other MOF collections.



Acknowledgments

David Hopkinson

Dan C. Sorescu

Surya Tiwari

Paul Boone

Patrick Muldoon

Clare Martin

This work was performed in support of the U.S. Department of Energy's Fossil Energy and Carbon Management's and executed through the National Energy Technology Laboratory (NETL) Research & Innovation Center