Project ID: ST144

HyMARC Seedling:
Optimized Hydrogen Adsorbents via Machine Learning and Crystal Engineering

Alauddin Ahmed,[1] Suresh Kuthuru,[2] Darpan Aulakh,[2] Justin Purewal,[3] Antek Wong-Foy,[2] Adam Matzger,[2] Mike Veenstra,[3] and **Don Siegel**[1] (PI)

[1]Mechanical Engineering Department and [2]Department of Chemistry, University of Michigan

[3]Ford Motor Company

DOE Annual Merit Review, May 1, 2019, Washington, DC

Overview

Timeline and Budget

Project Start Date: Sept. 1st, 2017
Project End Date: Dec. 31st, 2020

Total Project Budget: $1,047,000

Federal Share:

UM: $807,000
Ford: $192,000
Total: $999,000

$250,000 (Y1)
$398,000 (Y2)
$351,000 (Y3)

Cost Share: $48,000 (Ford)

Total Funds Spent:* ~$300,000

*Estimated as of 3/30/19

Barriers

Barriers addressed

- Volumetric Density
- Gravimetric Density

Partners

Interactions/collaborations:

Ford Motor Company, Hydrogen Storage Engineering Center of Excellence (HSECoE)

Project lead:

Don Siegel, University of Michigan

Background (1)

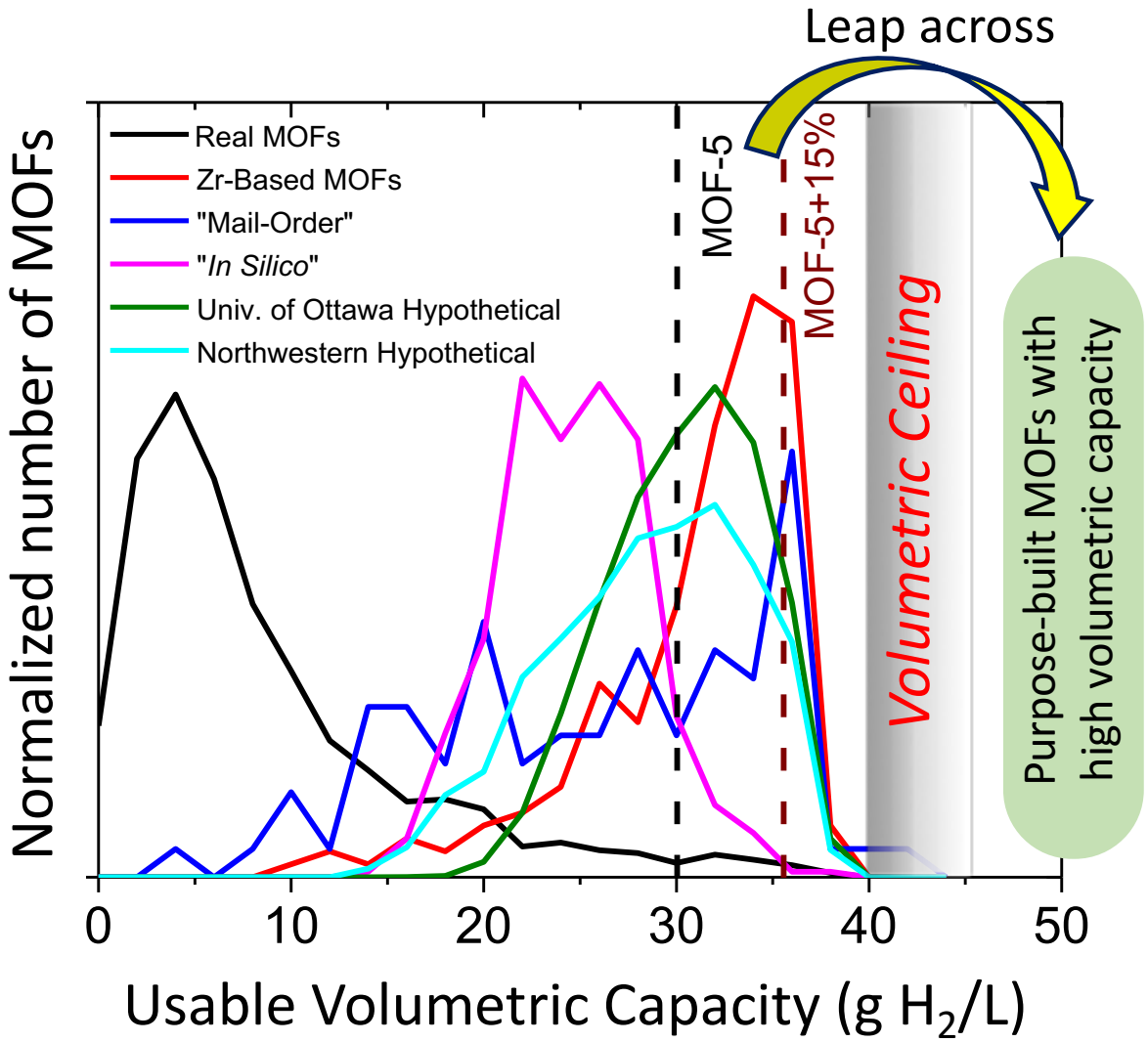
- A high-capacity, low-cost method for storing hydrogen remains one of the primary barriers to the widespread commercialization of fuel cell vehicles
- Storage via adsorption in MOFs is promising due to their fast kinetics, reversibility, and tunable properties
- A viable adsorbent must exhibit a high intrinsic (i.e., materials level) H_2 capacity, and pack in a dense fashion at the system level
  - Our prior screening revealed that no known MOF exhibits a usable volumetric capacity exceeding 40 g H_2/L (assuming a pressure swing between 100 and 5 bar at 77 K)
  - Analysis by the HSECoE has shown that inefficient materials packing can result in capacity reductions of more than 60% compared to the single-crystal level. These inefficiencies can negate improvements in volumetric performance achieved at the materials level
  - **This project addresses both of these challenges**

Project goal: Overcome volumetric limitations associated with physisorptive hydrogen storage at both the materials and systems level in metal-organic frameworks (MOFs)

Relevance (1)

Screening of ~500,000 MOFs reveals that essentially no compounds exceed 40 g/L usable capacity

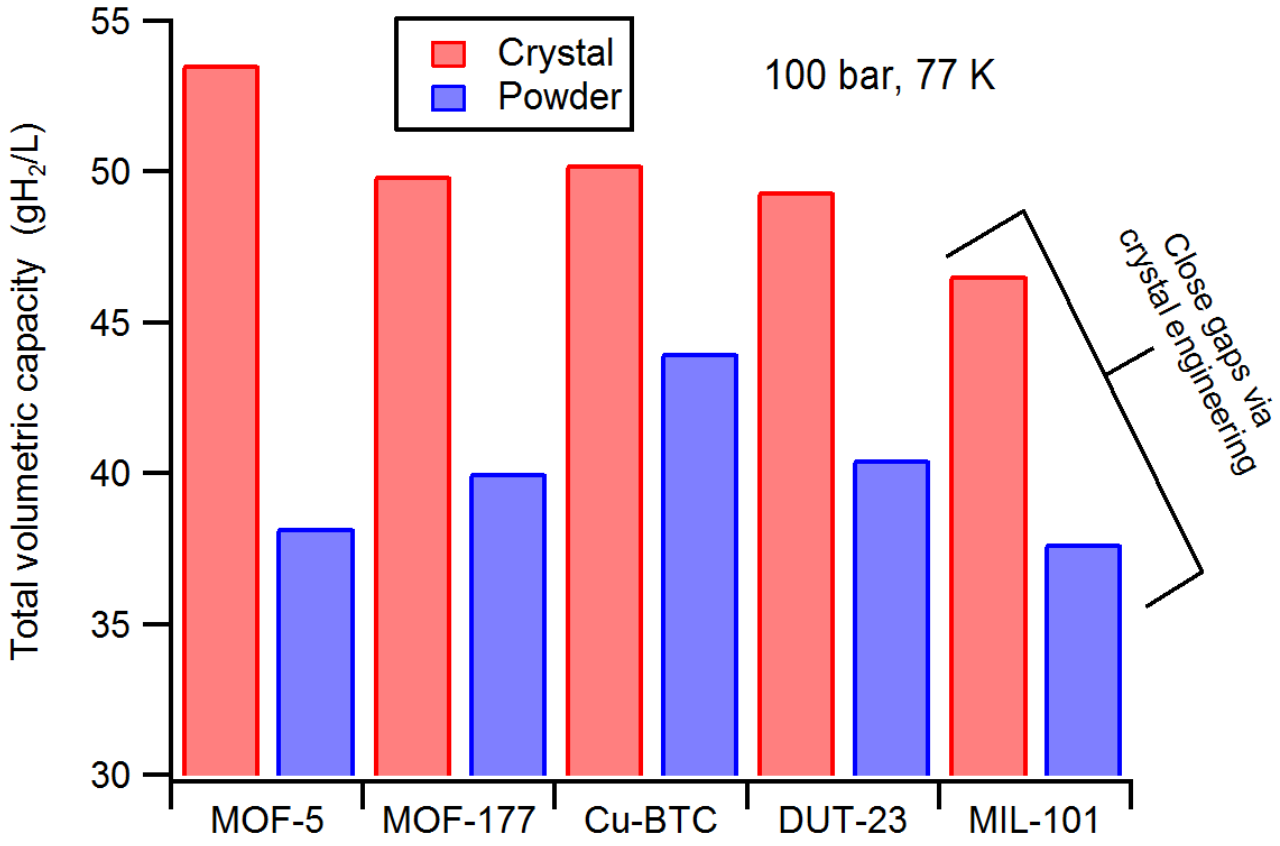
→ New MOFs needed to break through volumetric ceiling



Relevance (2)

Packing inefficiencies result in significant volumetric penalties in adsorptive hydrogen storage systems

→ Increase packing density via crystal engineering



Data courtesy of Justin Purewal, Ford Motor Company

Relevance (3)

Objective 1: Apply machine learning techniques to identify, design, and demonstrate high-capacity MOFs

- Demonstrate usable volumetric capacities exceeding 50 H_2 g/L (single-crystal/pressure swing)
- No compromise to gravimetric capacity, kinetic performance, or reversibility
- If successful, these compounds will set a new high-water mark for H_2 density in adsorbents at cryogenic conditions

Objective 2: Control MOF crystal morphology and crystallite size distribution to increase packing density

- Increase packing density of target high capacity MOF by at least 30% (compared to its powder tap density)
- Do so with less than 15% decrease in gravimetric performance

Milestones

Project has met all milestones through Q7

Milestone Summary Table						
Recipient Name:	University of Michigan (PI: Siegel)					
Project Title:	Optimized Hydrogen Adsorbents via Machine Learning and Crystal Engineering					
Task No.	**Task Title**	**Milestone Type**	**Milestone Description**	**Milestone Verification Process**	**Due Date (Month)**	**Status**
1.0	**MOF Discovery Guided by Machine Learning (ML)**					
1.1	MOF performance from scratch	Milestone	Demonstrate ability to predict usable capacity of an arbitrary MOF to within 85% of GCMC capacity using only crystal structure as input	Comparison of machine learning prediction with GCMC calculation	3	**Complete**
1.2	Structure-performance correlations	Milestone	Correlate MOF geometric properties with capacity	Random forest or SVM analysis analysis of MOF properties	6	**Complete**
1.3	MOF reverse engineering	Go/No-Go	Identify ranges for 4 MOF crystallographic properties (surface area, density, pore volume, & porosity) consistent with usable volumetric capacity of at least 40 g/L and usable gravimetric capacity of at least 7 wt. % (assuming an isothermal pressure swing between 100 and 5 bar at 77 K) based on single crystal density. Demonstrate that the identified ranges are within the realm of possibility for the development of new MOFs, and thus provide a pathway for meeting the DOE storage targets.	Random forest or SVM analysis of MOF properties and direct GCMC simulation	12	**Passed**
1.4	Data Dissemination	Milestone	Make list of structures available to HyMARC team and to general public via web download.	Data access confirmed by HyMARC partners	18	**Complete**
1.5	Validate ML predictions	Milestone	Use GCMC to validate ML predictions of highest capacity MOFs on a pressure swing and temperature+pressure swing basis. Attempt to synthesize 1-2 of the most promising MOF candidates. Assess surface areas; if within 85% of theoretical value perform PCT capacity measurements	GCMC calculations, BET surface area measurements, and PCT measurements	21	**Complete.** (Additional experiments on-going)
2.0	**Enhanced MOF Packing Density via Crystal Engineering**					
2.1	Morphological engineering	Milestone	Identify at least 2 additives capable of controlling morphology from cubes to octahedra	Optical microscopy or SEM measurements	9	**Complete**
2.1	Morphological engineering	Go/No-Go	Demonstrate an improvement in either A) a MOF with a single crystal volumetric capacity greater than 39 g/L usable capacity measured at 77 K, and 5-100 bar pressure (i.e., a 10% increase over the current state-of-the-art NU-100) through Machine Learning-directed material development, OR B) a 15% increase in tap density through crystal engineering methods for a specific MOF compared to its non-optimized powder, with a minimal loss in surface area.	Density measurements, surface area measurements, PCT measurements	24	**In progress**
2.2	Particle size control	Milestone	Determine if particle size influences packing efficiency by more than 10% for particles whose size varies by more than an order of magnitude	Void fraction measurements	15	**Complete**

Approach

MOF Database

Compiled a MOF database of ~500,000 compounds

~100,000 MOFs assessed for temperature + pressure swing storage

~100,000 MOFs assessed for pressure swing storage

Source	Available in database	Zero surface area	H_2 capacity evaluated empirically	H_2 capacity evaluated with GCMC
UM+CoRE+CSD17 (RM)	15,235	2,950	12,285	12,799
Mail-Order MOFs (MO)	112	4	108	112
In Silico MOFs (IS)	2,816	154	2,662	466
In Silico Surface MOFs (ISS)	8, 885	283	8,602	1,058
MOF-74 Analogs (M74)	61	0	61	61
ToBaCCo (TB)	13,512	214	13,298	2,854
Zr-MOFs (ZR)	204	0	204	204
NW Hypothetical MOFs (NW)	137,000	30,160	106,840	20,156
UO Hypothetical MOFs (UO)	324,500	32,993	291,507	61,247
In-house synthesized via hypothetical design	18	0	18	5
Total	**493,458**	**66,758**	**426,700**	**98,962**

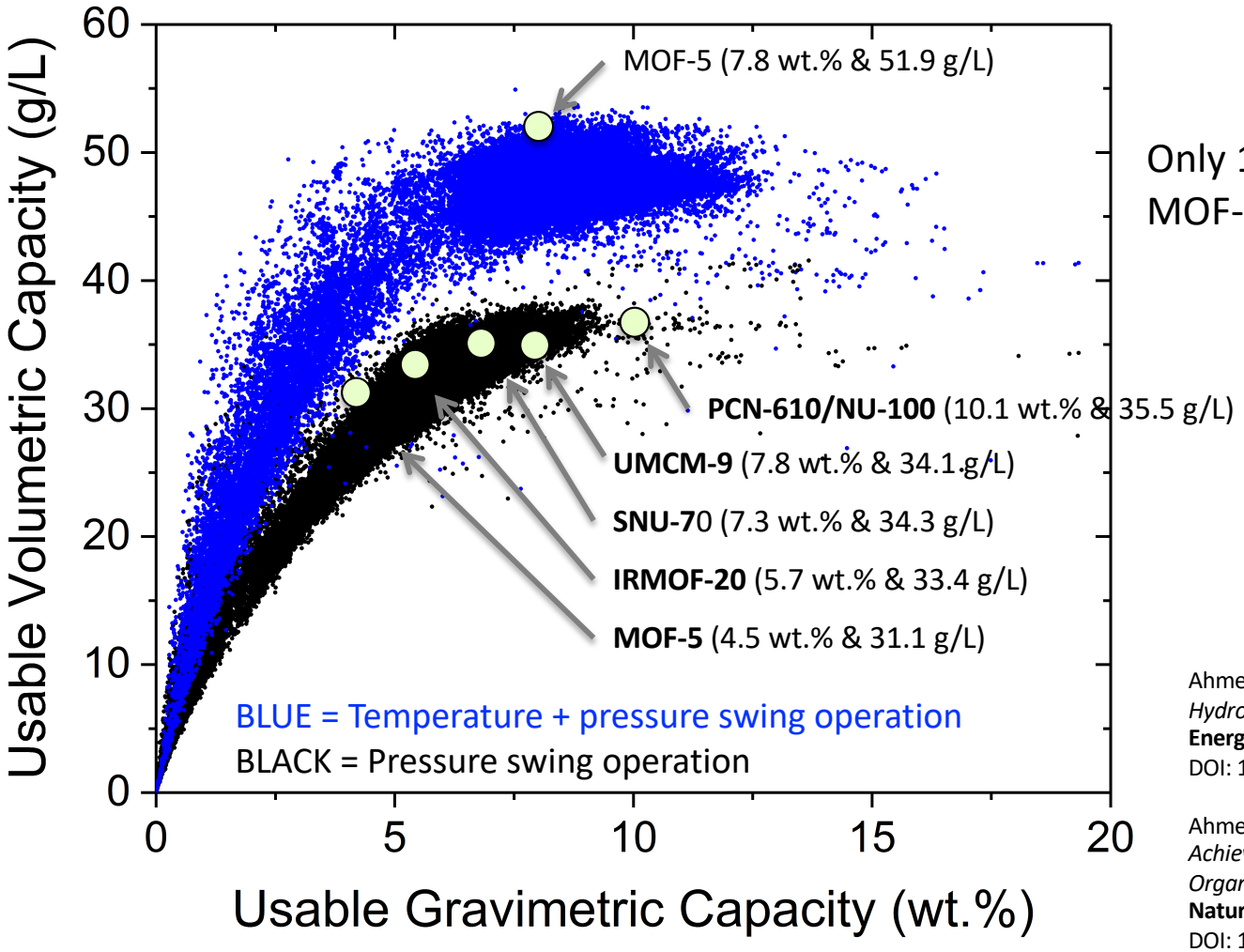
RM: (a) **UM:** J.Goldsmith, A. G. Wong-Foy, M. J. Cafarella, and D. J. Siegel, *Chem. Mater.,* 25 , 3373–3382 (2013); (b) **CoRE:** Y. G. Chung, *et al*., *Chem. Mater*., 26, 6185–6192 (2014); (c) **CSD17**: P. Z. Moghadam et al., *Chem. Mater.,* 29, 2618–2625 (2017).
MO: R. L. Martin, L.-C. Lin, K. Jariwala, B. Smit, M. Haranczyk, *J. Phys. Chem. C* 117, 12159-12167 (2013);
IS: Y. Bao, R. L. Martin, M. Haranczyk, M. W. Deem, *J. Phys. Chem. C* 119, 186-195 (2015).
ISS: Y. Bao, R. L. Martin, C. M. Simon, M. Haranczyk, B. Smit, M. W. Deem, *Phys. Chem. Chem. Phys.,* 17, 11962-11973 (2015).
M74: M. Witman, S. Ling, S. Anderson, L. Tong, K.C. Stylianou, B. Slater, B. Smit, M. Haranczyk, *Chem. Sci*., 7, 6263-6272 (2016).
TB: Y. J. Colón, D. A. Gómez-Gualdrón, and R. Q. Snurr, *Cryst. Growth Des.,* 17, 5801–5810 (2017).
ZR: D. A. Gómez-Gualdrón, O.V. Gutov, V. Krungleviciute, B. Borah, J. E. Mondloch, J. T. Hupp, T. Yildirim, O.K. Farha, R.Q. Snurr, *Chem. Mater.* 26, 5632-5639 (2014).
NW: C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp, R. Q. Snurr, *Nat. Chem.* 4, 83–89 (2012).
UO: M. Z. Aghaji, M. Fernandez, P. G. Boyd, T. D. Daff, and T. K. Woo, *Eur. J. Inorg. Chem.,* 2016, 4505–4511 (2016).

High-Throughput Screening

Predicted usable H_2 capacities for PS and TPS conditions

Pressure swing: P_{max}= 100 bar to P_{min}= 5 bar at 77 K

Temp+pressure swing: T_{min}= 77 K, P_{max}= 100 bar to T_{max}= 160 K, P_{min}= 5 bar



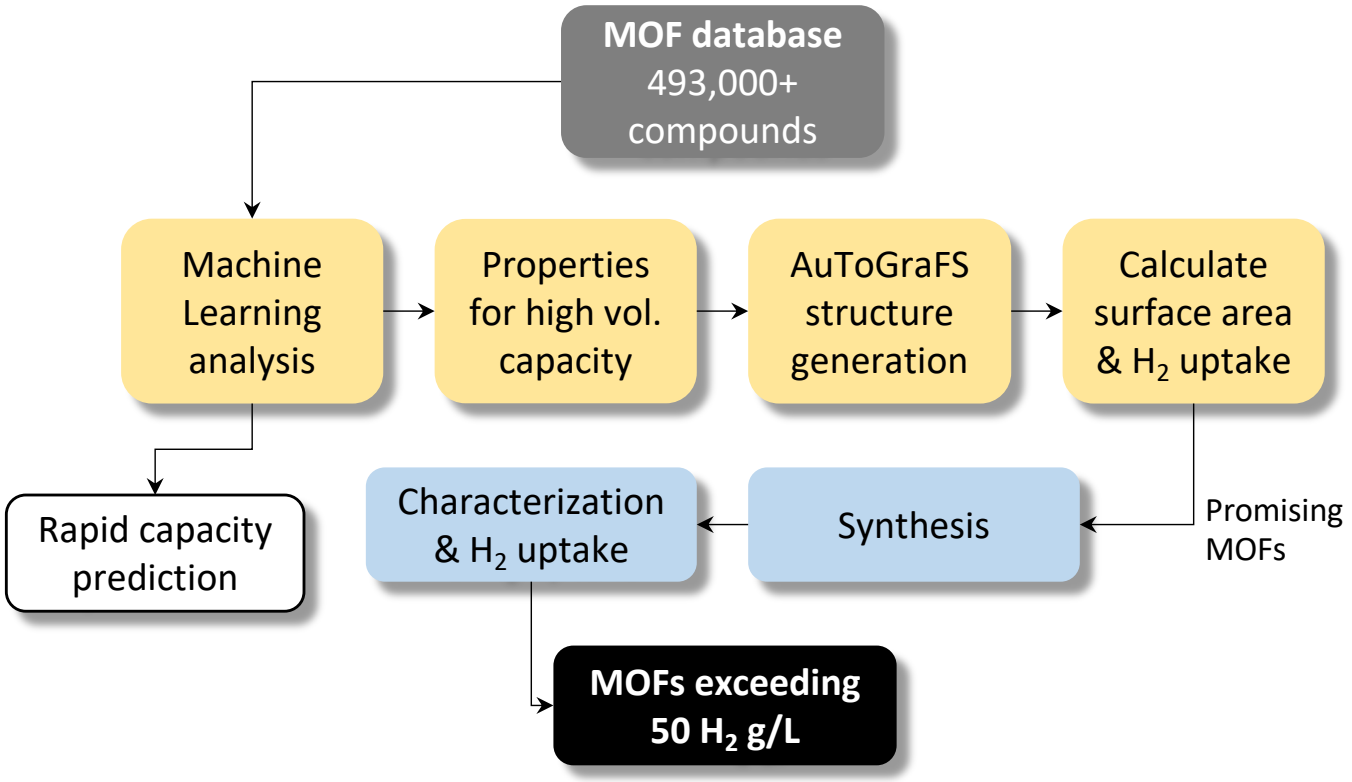
Only 180 MOFs surpass MOF-5 under TPS conditions

Ahmed et al., *Balancing Gravimetric and Volumetric Hydrogen Density in MOFs,*
Energy & Environmental Science, 10, 2459 (2017)
DOI: 10.1039/C7EE02477K

Ahmed et al., *Exceptional Hydrogen Storage Achieved by Screening Nearly Half a Million Metal-Organic Frameworks,*
Nature Communications, in press (2019)
DOI: 10.1038/s41467-019-09365-w

Concept

Machine learning will be used to guide the development of MOFs with high volumetric H_2 capacities



Crystal Engineering

- Packing of congruent convex objects indicates that particle morphology and the size distribution are key factors in determining packing efficiency
- We shall vary these properties systematically, leveraging advances in colloid science for the controlled growth of MOFs with various shapes and sizes

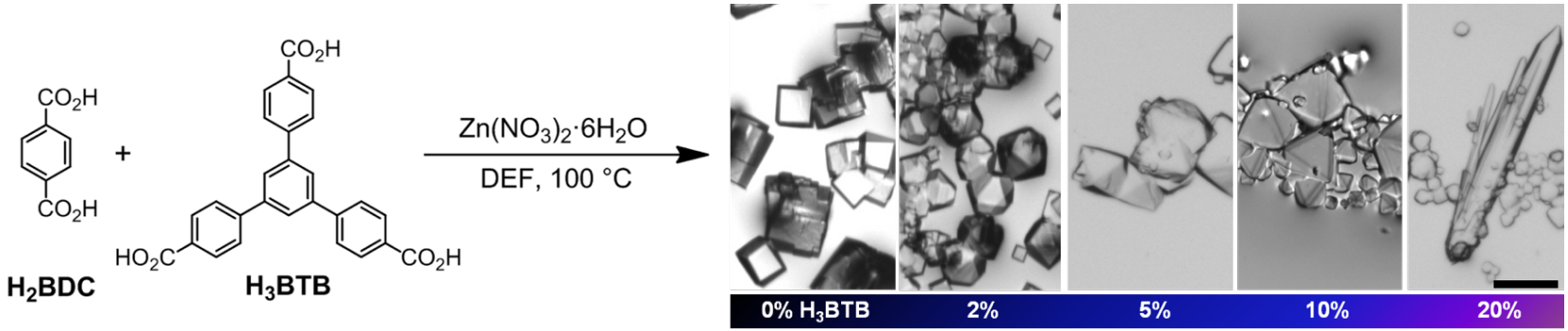
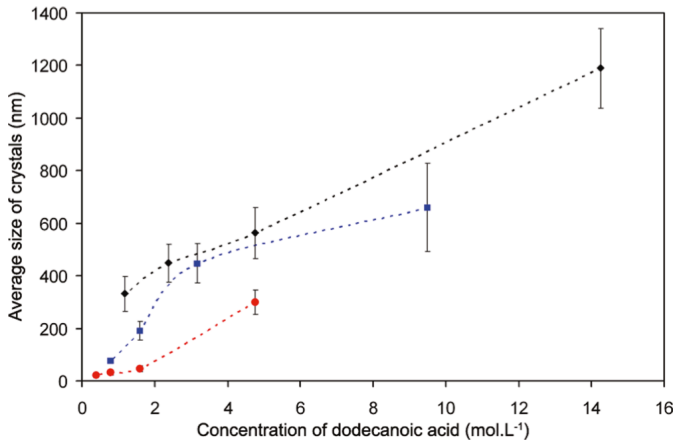


Fig. 1: Synthesis of octahedral-shaped MOF-5 crystals by addition of H_3BTB in the reaction mixture of H_2BDC and $Zn(NO_3)_2 \cdot 6H_2O$. Photographs show the dependence of crystal morphology on the percentage of H_3BTB (scale bar: 100 μm). Another phase (needle shaped UMCM-1) appears at 10 mol% H_3BTB. From Matzger et al., JACS (2011) **133,** 20138

Fig. 2: Average size of HKUST-1 crystals as a function of dodecanoic acid concentration taken at longer and longer times. Colors represent different concentration of dodecanoic acid. From Diring, et al., Chem. Mater., (2010) **22**, 4531



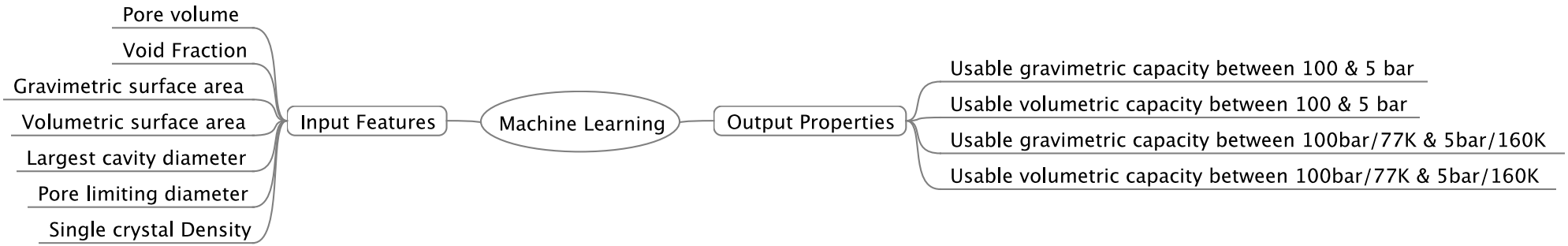
Accomplishments and Progress

Machine Learning Approach

ML was used to make capacity predictions from crystallographic features

7 Crystallographic features calculated via Zeo++

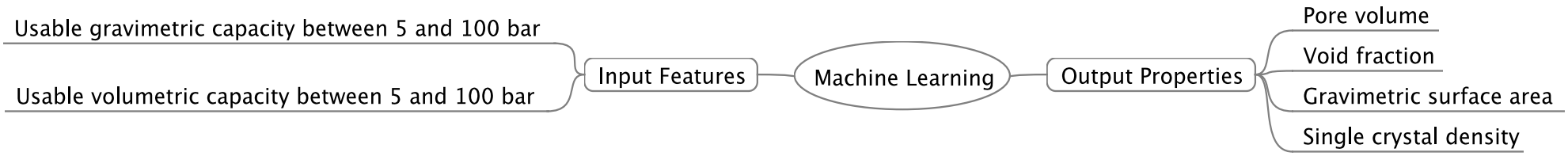
H_2 storage capacities at 4 conditions calculated via GCMC



The reverse process was used to predict of crystallographic properties consistent with a target capacity

Usable H_2 storage capacities at 77 K for PS between 5 & 100 bar calculated via GCMC

4 crystallographic features calculated via Zeo++



*Zeo++: T.F. Willems, C.H. Rycroft, M. Kazi, J.C. Meza, and M. Haranczyk, Microporous and Mesoporous Materials, 149 ,134-141 (2012).

ML Methods Tested

12 supervised learning methods from 5 different categories were used

Decision Trees (DT)

L. Breiman, J. Friedman, R. Olshen, and C. Stone. Classification and Regression Trees. Wadsworth, Belmont, CA, 1984.

Ensemble Methods

Random Forest (RF)

L. Breiman, "Random Forests", Machine Learning, 45(1), 5-32, 2001.

Bagging with DT; Bagging with RF (Bagging)

L. Breiman, "Bagging predictors", Machine Learning, 24(2), 123-140, 1996.

Boosted DT; Ada Boost with RF (Ada Boost)

H. Drucker. "Improving Regressors using Boosting Techniques", 1997

Gradient Boosting

J. Friedman, Greedy Function Approximation: A Gradient Boosting Machine, The Annals of Statistics, Vol. 29, No. 5, 2001.

Extremely Randomized Trees

P. Geurts, D. Ernst., and L. Wehenkel, "Extremely randomized trees", Machine Learning, 63(1), 3-42, 2006.

K-Nearest Neighbors (K-NN)

N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression". The American Statistician. 46(3), 175–185, 1992.

Support Vector Machine (SVM)

A. J. Smola, B. Schölkopf, "A Tutorial on Support Vector Regression", Statistics and Computing archive, 14(3), 199-222, 2004.

Linear Regression; Ridge Regression (Generalized Linear Model)

T. Hastie, R. Tibshirani and J. Friedman. Elements of Statistical Learning Ed. 2, Springer, 2009.

Machine Learning Software & Code: Scikit-learn, R, & in-house code

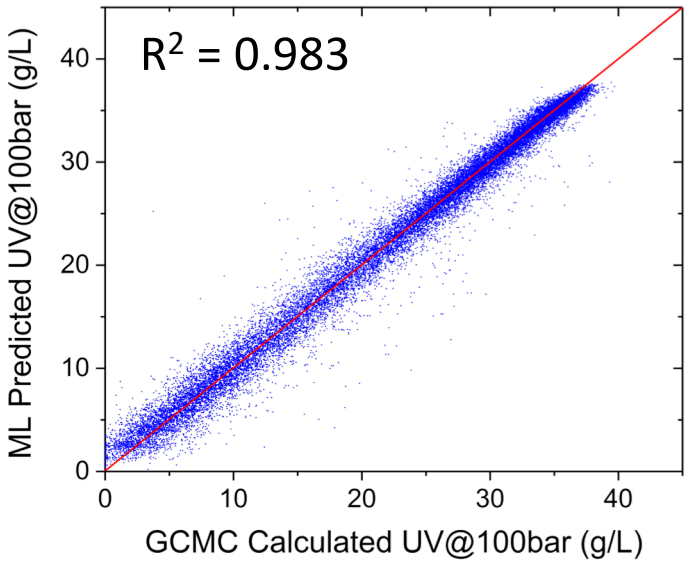
Hyperparameter Optimization Method: Grid search cross validation method

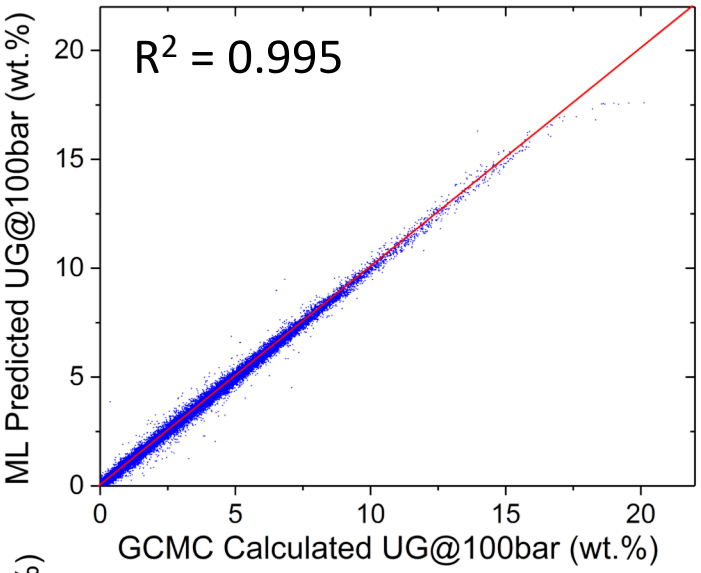
Training Set: 74, 221; Test Set: 24,741; Unseen Data Set: 394,496

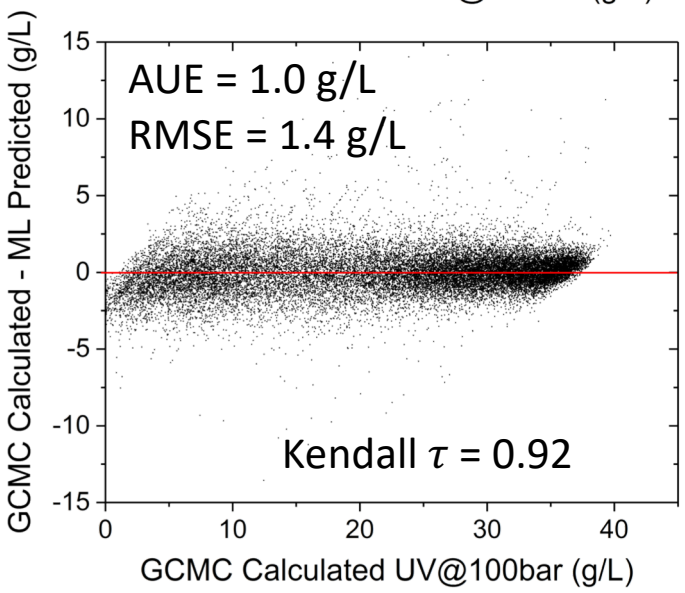
- Scikit-learn: Pedregosa *et al.*, Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825-2830, 2011.
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

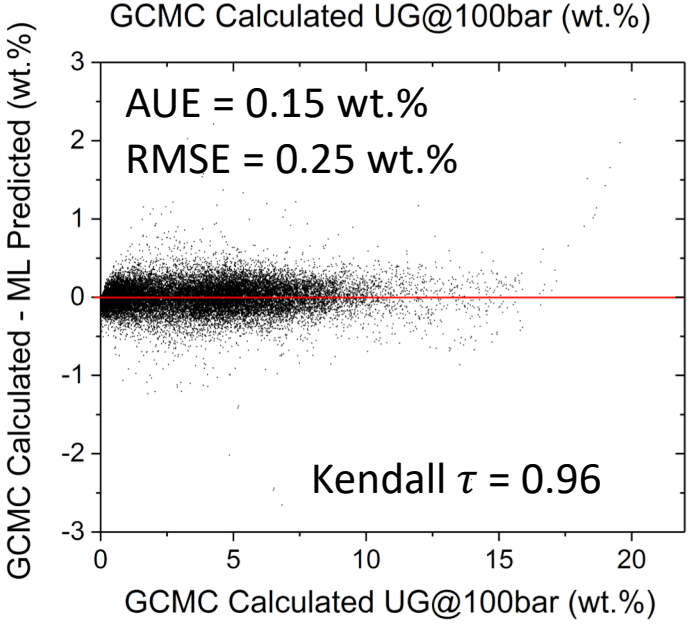
ML Predictions of Capacities

Comparison between Extremely Randomized Trees ML prediction and GCMC





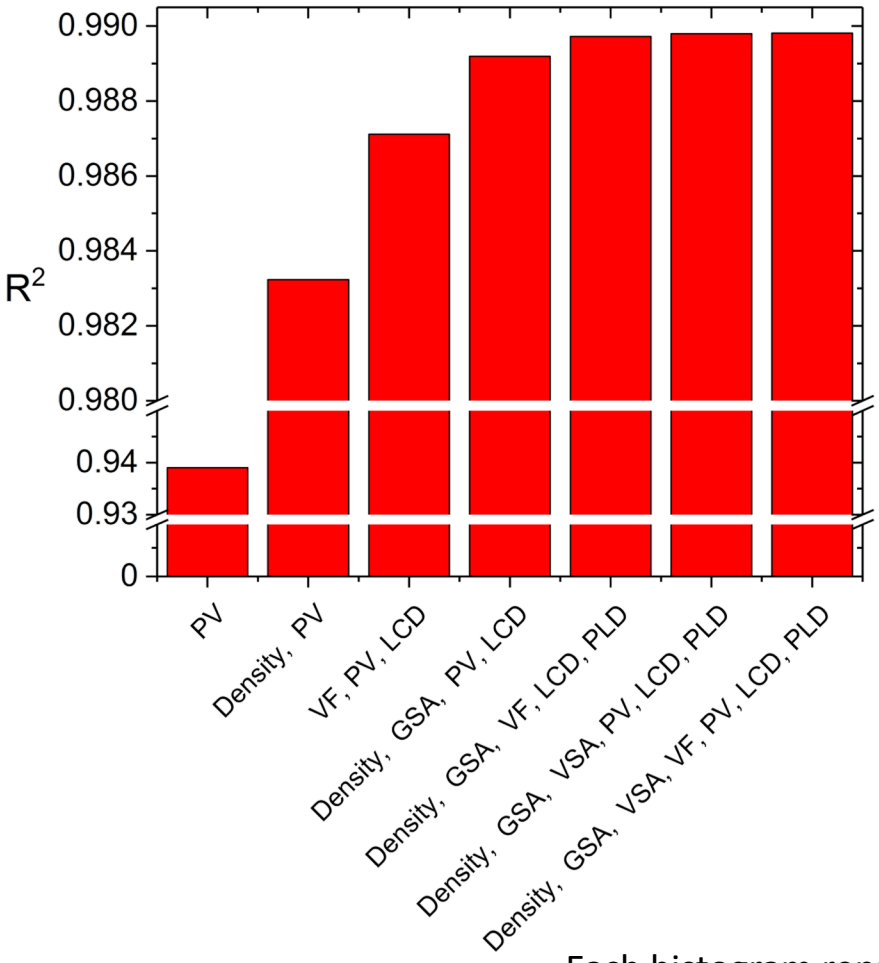




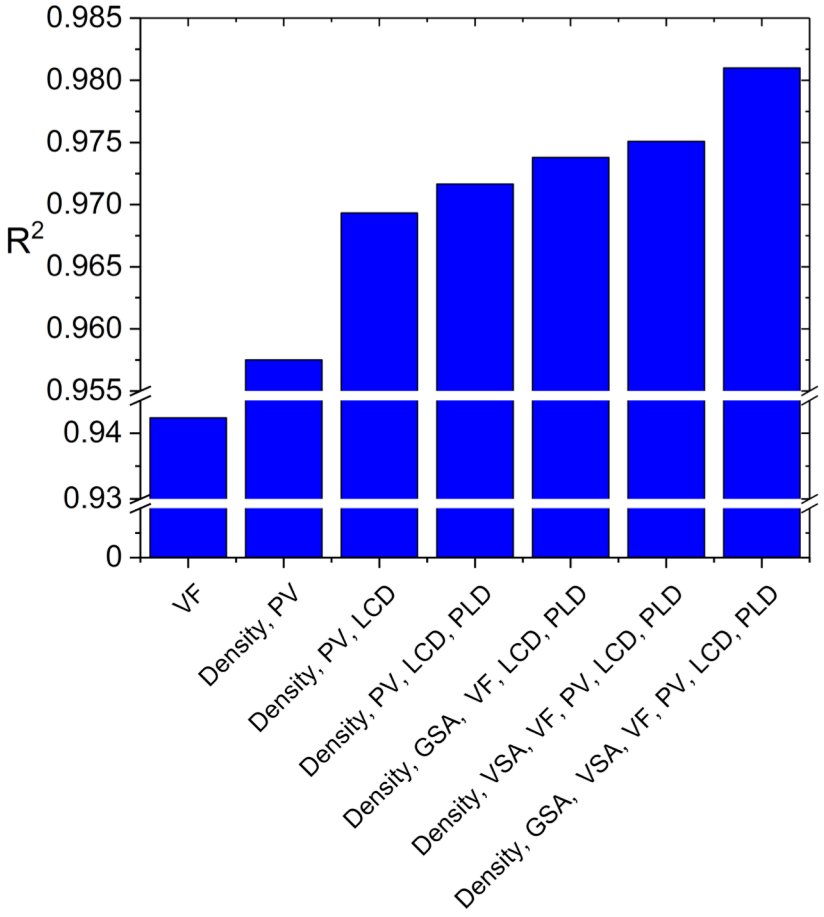
Structure-Property Correlations

ML models were developed for all possible combinations of crystallographic features to identify the optimal feature set

Usable Gravimetric



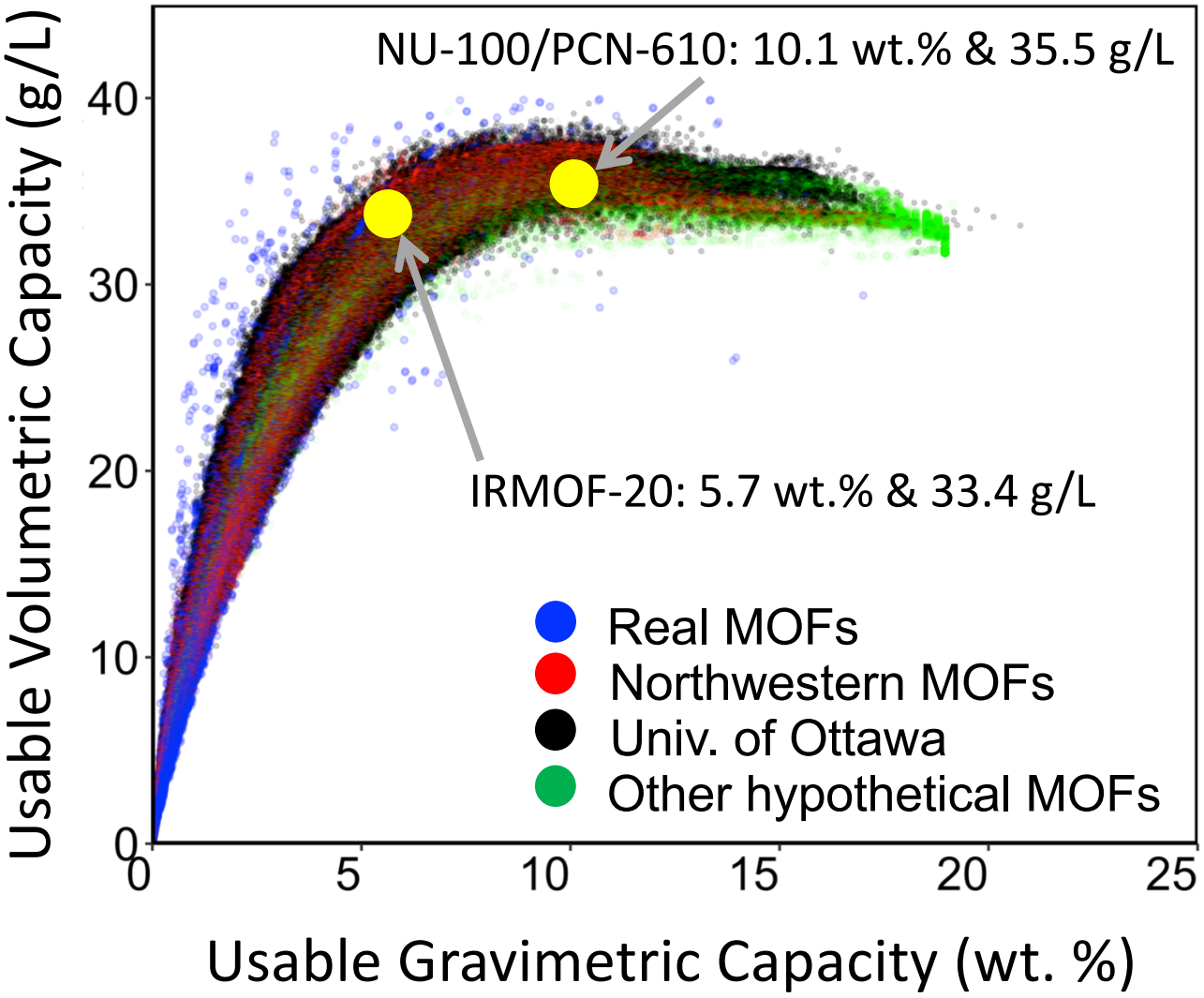
Usable Volumetric



Each histogram represents the highest R^2 value among all possible combinations of a given number of features.

H_2 Storage in 500,000 MOFs

ML reveals 69,363 MOFs that can potentially out-perform IRMOF-20



Go/No-Go

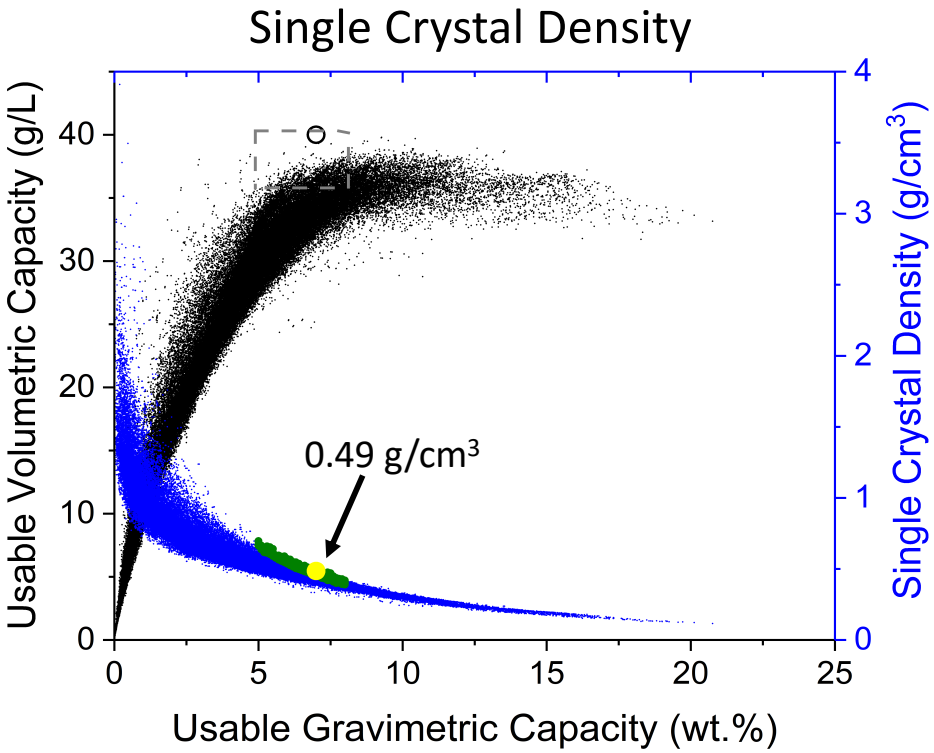
The Y1 Go/No-Go milestone was successfully met

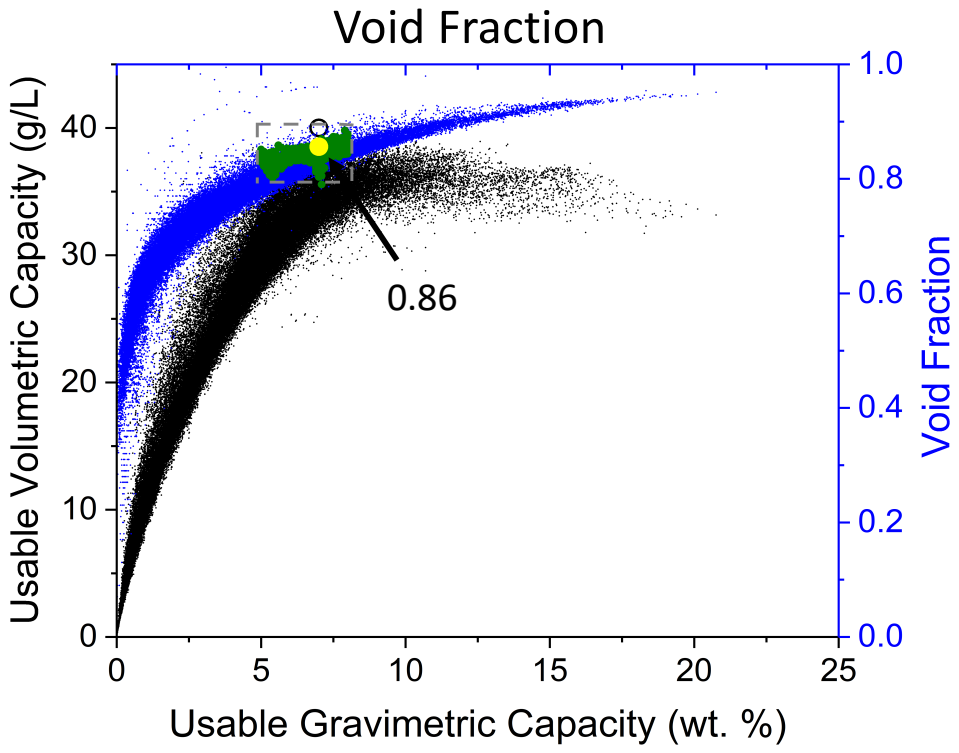
Task No.	Task Title	Milestone Type	Milestone Description	Milestone Verification Process	Due (Date)	Status
1.3	MOF reverse engineering	Go/No-Go	Identify ranges for 4 MOF crystallographic properties (surface area, density, pore volume, & porosity) consistent with usable volumetric capacity of at least 40 g/L and usable gravimetric capacity of at least 7 wt. % (assuming an isothermal pressure swing between 100 and 5 bar at 77 K) based on single crystal density. Demonstrate that the identified ranges are within the realm of possibility for the development of new MOFs, and thus provide a pathway for meeting the DOE storage targets.	Random forest or SVM analysis of MOF properties and direct GCMC simulation	8/31/18	**Passed**

Crystallographic Property	ML Prediction
Density (g/cm^3)	0.49 ± 0.01
Gravimetric Surface Area (m^2/g)	5222 ± 402
Pore Volume (cm^3/g)	1.74 ± 0.03
Void Fraction	0.86 ± 0.02

Go/No-Go

ML models were trained 'in reverse' to predict the crystallographic propoperties that correspond to a specified usable capacity





Range of capacities: 5-8 wt. % and 36-40 g/L

Crystallographic properties corresponding to UG & UV range (dashed region)

ML input capacities: 7 wt.% & 40 g/L

Crystallographic property corresponding to 7 wt.% & 40 g/L

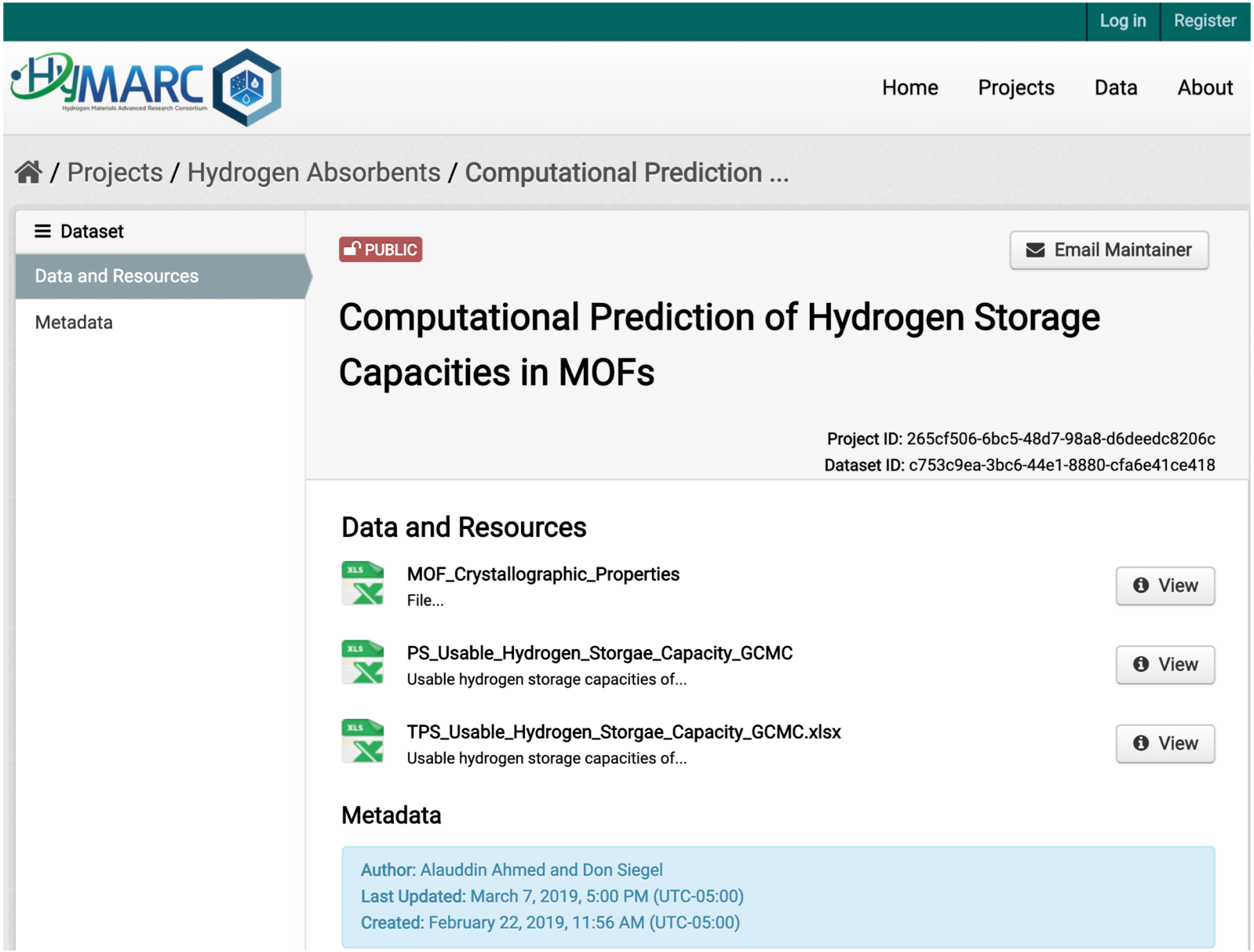
Milestone 1.4

Data generated from our prior high-throughput screening studies have been uploaded to the HyMARC data hub

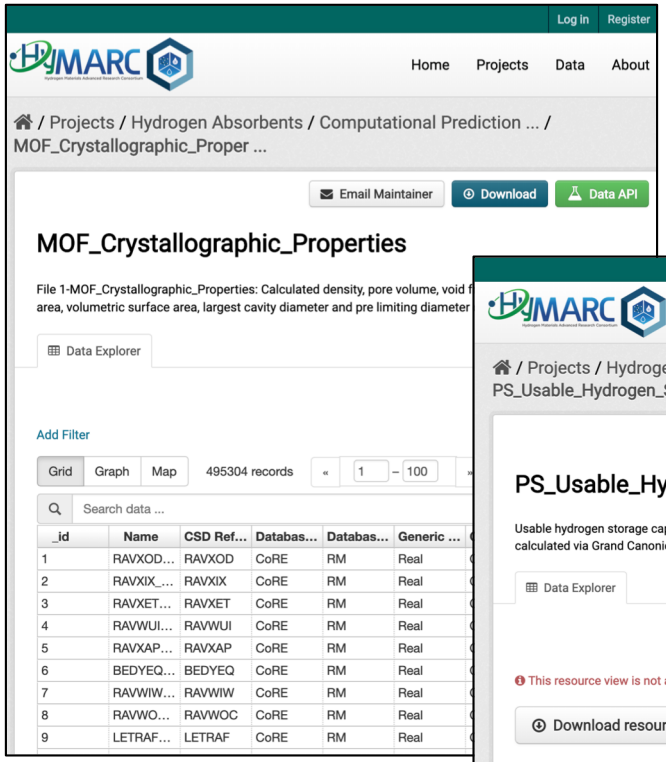
Task No.	Task Title	Milestone Description	Milestone Verification Process	Due Date (Month)	Status
1.4	Data Dissemination	Make list of structures available to HyMARC team and to general public via web download.	Data access confirmed by HyMARC partners	18	**Complete**

Data Dissemination

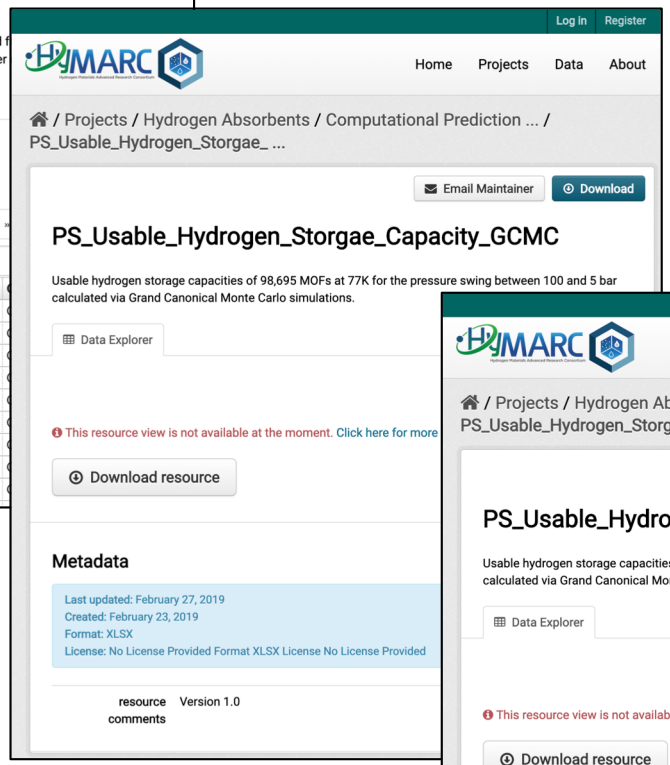
https://datahub.hymarc.org



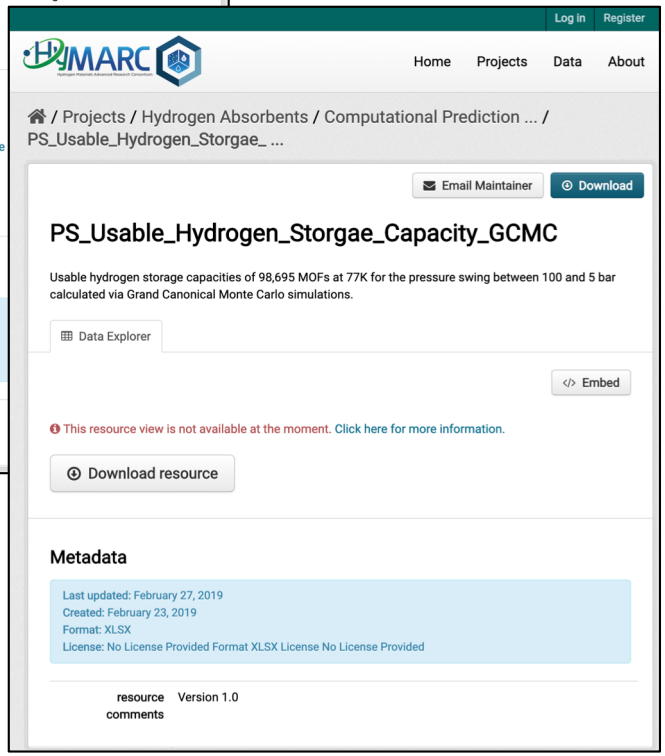
Data Dissemination



Crystallographic properties of 495,305 MOFs



Usable pressure swing capacities of 98,695 MOFs



Usable temp + pressure swing capacities of 98,695 MOFs

Milestone 1.5

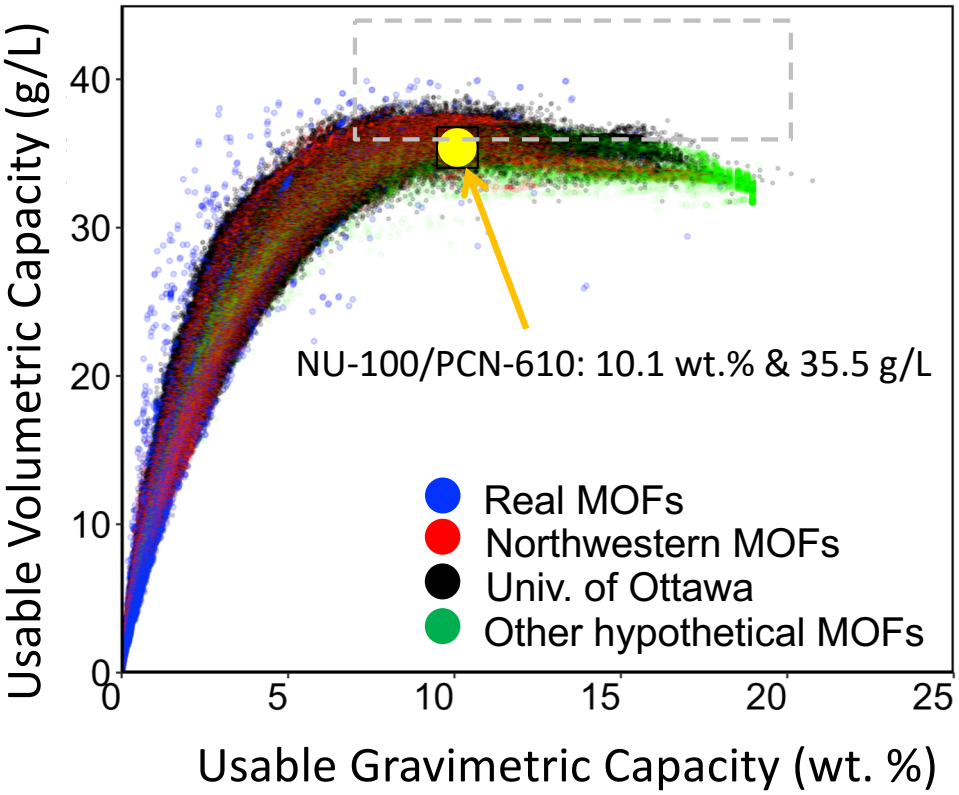
Milestone has been met.

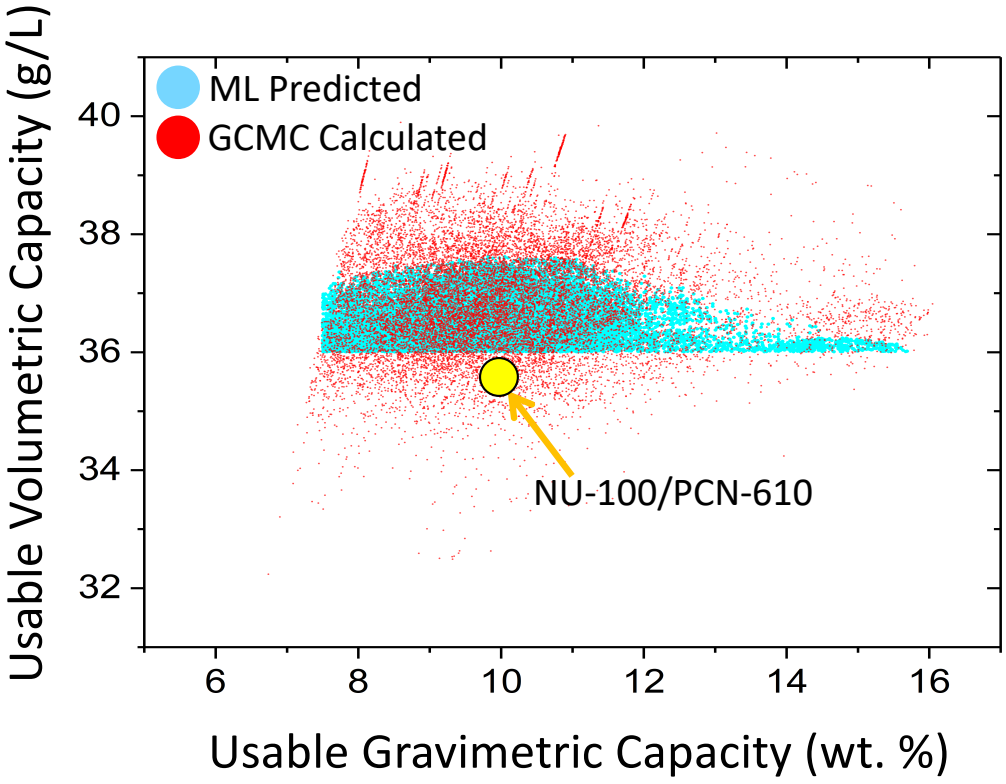
Additional experimental work is on-going to realize high-capacity MOFs for temperature + pressure swing operation

Task No.	Task Title	Milestone Description	Milestone Verification Process	Due Date (Month)	Status
1.5	Validate ML predictions	Use GCMC to validate ML predictions of highest capacity MOFs on a pressure swing and temperature + pressure swing basis. Attempt to synthesize 1-2 of the most promising MOF candidates. Assess surface areas; if within 85% of theoretical value perform PCT capacity measurements	GCMC calculations, BET surface area measurements, and PCT measurements	21	**Complete.** (Additional experiments on-going)

Pressure Swing

Identified 6,336 MOFs (out of ~500,000) using ML that surpass the usable capacity of NU-100/PCN-610 under pressure swing conditions (100 → 5 bar at 77K.





MOFs predicted by ML to be the highest capacity were validated using more accurate GCMC calculations

Pressure Swing

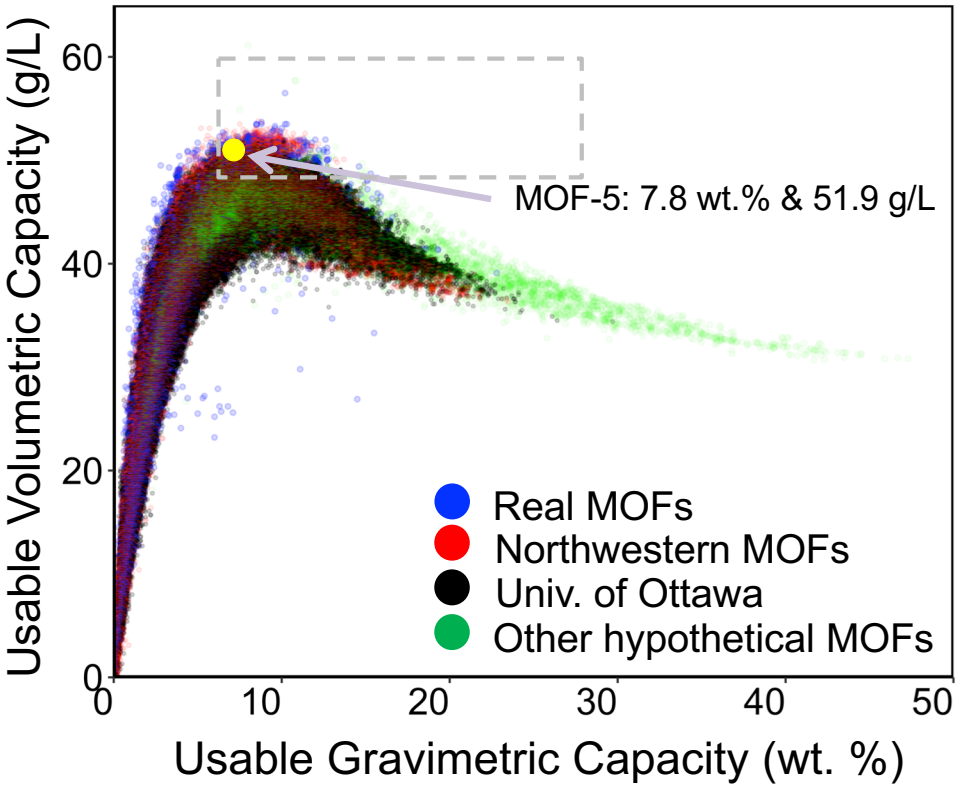
List of highest-capacity MOFs and their respective capacities as predicted by ML and subsequent GCMC calculations

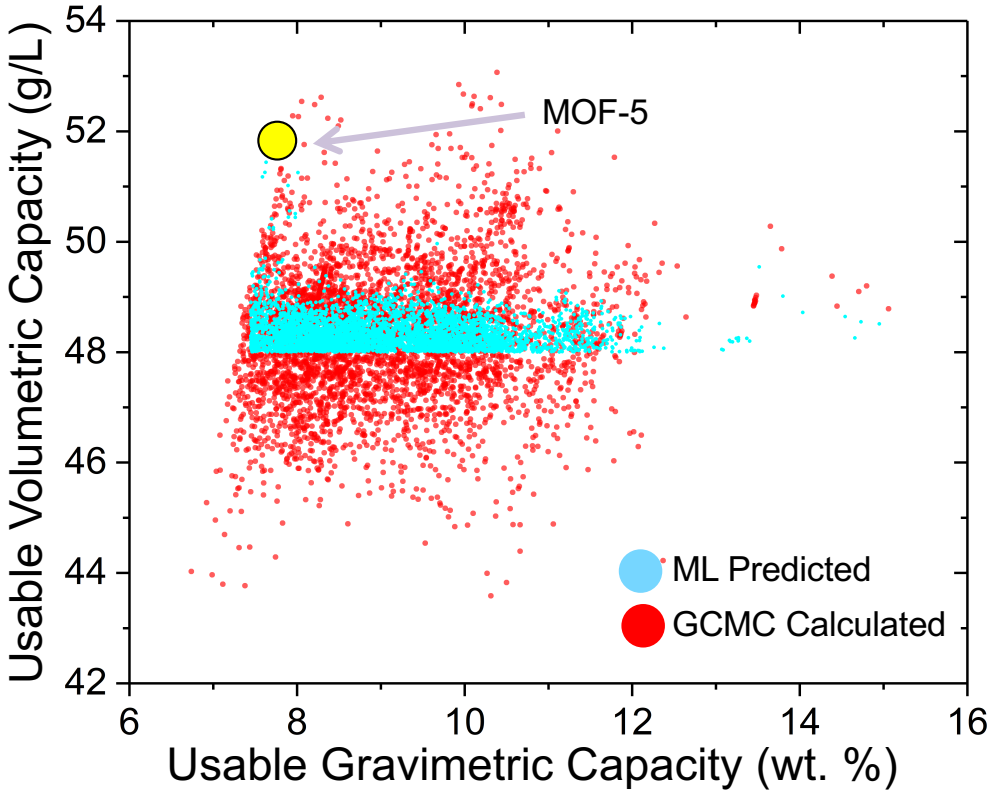
MOF Name	Source	Density (g/cm^3)	Gravimetric Surface Area (m^2/g)	Volumetric Surface Area (m^2/cm^3)	Void Fraction	Pore Volume (cm^3/g)	Largest Cavity Diameter (Å)	Pore Limiting Diameter (Å)	Usable Grav. Capacity (wt. %)			Usable Vol. Capacity (g/L)		
									GCMC	ML	GCMC-ML	GCMC	ML	GCMC-ML
mof_7642	ToBaCCo	0.30	5561	1695	0.89	2.93	12.8	11.8	11.1	10.3	0.8	40.5	37.4	3.0
mof_7690	ToBaCCo	0.30	5715	1706	0.89	2.98	12.8	12.0	11.3	10.4	0.9	40.3	37.3	3.0
mof_7594	ToBaCCo	0.40	5070	2031	0.86	2.15	11.2	9.7	8.6	7.9	0.7	39.9	37.0	2.9
mof_7210	ToBaCCo	0.29	5936	1730	0.89	3.04	13.4	11.7	11.4	10.5	0.9	39.8	37.1	2.8
mof_7738	ToBaCCo	0.25	6054	1502	0.90	3.64	14.5	13.5	13.0	12.0	1.0	39.7	37.0	2.7
hypotheticalMOF_5045702_i_1_j_24_k_20_m_2	NW	0.31	5926	1820	0.88	2.87	16.0	11.0	10.9	10.1	0.8	39.7	37.2	2.5
str_m3_o19_o19_f0_nbo.sym.1.out	UO	0.31	5073	1583	0.90	2.88	17.7	12.9	10.8	10.1	0.7	39.7	37.1	2.6
hypotheticalMOF_5037315_i_1_j_20_k_12_m_1	NW	0.31	5818	1787	0.88	2.86	16.0	11.0	10.9	10.0	0.9	39.7	37.0	2.6
hypotheticalMOF_5037467_i_1_j_20_k_12_m_8	NW	0.31	5860	1800	0.88	2.85	16.0	11.0	10.9	10.0	0.9	39.7	37.0	2.7
str_m3_o5_o20_f0_nbo.sym.1.out	UO	0.39	4772	1882	0.87	2.22	14.1	9.6	8.7	8.1	0.7	39.7	37.2	2.5
hypotheticalMOF_5037563_i_1_j_20_k_12_m_13	NW	0.31	5897	1811	0.88	2.87	16.1	11.0	10.9	10.1	0.8	39.7	37.2	2.5
hypotheticalMOF_5038404_i_1_j_20_k_20_m_15	NW	0.31	5870	1803	0.88	2.87	16.0	11.0	10.9	10.1	0.8	39.7	37.2	2.5
hypotheticalMOF_5037379_i_1_j_20_k_12_m_4	NW	0.31	5818	1787	0.88	2.86	16.0	11.0	10.9	10.0	0.8	39.6	37.0	2.6
hypotheticalMOF_5037407_i_1_j_20_k_12_m_5	NW	0.31	5818	1787	0.88	2.86	16.0	11.0	10.9	10.0	0.8	39.6	37.0	2.6
hypotheticalMOF_5037479_i_1_j_20_k_12_m_9	NW	0.31	5818	1787	0.88	2.86	16.0	11.0	10.9	10.0	0.8	39.6	37.0	2.6
hypotheticalMOF_5055561_i_1_j_28_k_20_m_11	NW	0.31	5874	1804	0.88	2.87	16.0	11.0	10.9	10.1	0.8	39.6	37.2	2.4
hypotheticalMOF_5037439_i_1_j_20_k_12_m_7	NW	0.31	5858	1799	0.88	2.85	16.0	11.0	10.9	10.0	0.9	39.6	37.0	2.6
hypotheticalMOF_5037499_i_1_j_20_k_12_m_10	NW	0.31	5854	1798	0.88	2.85	16.0	11.0	10.9	10.0	0.9	39.6	37.0	2.6
hypotheticalMOF_5037531_i_1_j_20_k_12_m_11	NW	0.31	5818	1787	0.88	2.86	16.0	11.0	10.9	10.0	0.8	39.6	37.0	2.6
hypotheticalMOF_5037523_i_1_j_20_k_12_m_11	NW	0.31	5857	1799	0.88	2.86	16.0	11.0	10.9	10.0	0.8	39.6	37.1	2.5
NU-100/PCN-610									10.1			35.5		
NU-100/PCN-610 + 10%									11.1			39.1		

GCMC and ML generally agree within 1 wt.% and 2-3 g/L

Temp + Pressure Swing

Identified only 20 MOFs that surpass the usable capacity of MOF-5 under temperature + pressure swing conditions between 100 bar/77 K and 5 bar/160 K





Temp + Pressure Swing

List of highest-capacity MOFs and their respective capacities as predicted by ML and subsequent GCMC calculations for 400,000 MOFs 'unseen' by ML

MOF Name	Source	Density (g/cm^3)	Gravimetric Surface Area (m^2/g)	Volumetric Surface Area (m^2/cm^3)	Void Fraction	Pore Volume (cm^3/g)	Largest Cavity Diameter (Å)	Pore Limiting Diameter (Å)	Usable Gravimetric Capacity (wt. %)			Usable Volumetric Capacity (g/L)		
									GCMC	ML	GCMC-ML	GCMC	ML	GCMC-ML
str_m1_o1_o11_f0_pcu.sym.102.out	UO	0.45	4352	1974	0.84	1.84	12.9	10.1	10.4	9.7	0.7	53.1	48.1	4.9
str_m1_o1_o11_f0_pcu.sym.117.out	UO	0.47	4162	1977	0.83	1.74	12.8	9.9	9.9	9.0	0.9	52.8	48.0	4.8
str_m1_o1_o11_f0_pcu.sym.121.out	UO	0.47	4263	2006	0.83	1.76	12.1	10.2	10.0	9.4	0.6	52.7	48.1	4.6
str_m1_o1_o11_f0_pcu.sym.13.out	UO	0.46	4326	2005	0.83	1.79	12.7	9.9	10.1	9.3	0.8	52.6	48.0	4.6
str_m1_o1_o11_f0_pcu.sym.159.out	UO	0.58	3703	2138	0.80	1.38	10.4	8.6	8.3	7.6	0.7	52.6	48.5	4.1
str_m1_o1_o11_f0_pcu.sym.200.out	UO	0.45	4359	1978	0.84	1.84	12.9	10.1	10.3	9.6	0.7	52.6	48.1	4.5
str_m1_o1_o11_f0_pcu.sym.212.out	UO	0.60	3417	2035	0.83	1.39	12.0	10.1	8.1	7.5	0.5	52.5	48.1	4.4
str_m1_o1_o11_f0_pcu.sym.51.out	UO	0.46	4330	2007	0.83	1.79	11.9	9.9	10.1	9.3	0.8	52.5	48.1	4.4
str_m1_o1_o11_f0_pcu.sym.71.out	UO	0.45	4436	1980	0.84	1.87	13.0	10.9	10.4	9.7	0.8	52.5	48.1	4.4
str_m1_o1_o11_f0_pcu.sym.89.out	UO	0.58	3507	2043	0.83	1.42	12.4	9.8	8.2	7.7	0.5	52.5	48.1	4.4
str_m1_o1_o17_f0_pcu.sym.1.out	UO	0.46	4283	1985	0.83	1.79	11.9	9.9	10.1	9.4	0.7	52.5	48.3	4.2
str_m1_o1_o17_f0_pcu.sym.104.out	UO	0.46	4439	2032	0.83	1.82	12.5	11.0	10.2	9.6	0.6	52.4	48.2	4.2
str_m1_o1_o17_f0_pcu.sym.129.out	UO	0.60	3585	2157	0.83	1.37	14.6	9.2	7.9	7.6	0.3	52.3	48.2	4.1
str_m1_o1_o17_f0_pcu.sym.132.out	UO	0.60	3438	2048	0.83	1.39	12.7	10.8	8.0	7.8	0.2	52.3	48.3	4.0
str_m1_o1_o17_f0_pcu.sym.28.out	UO	0.57	3732	2117	0.80	1.41	13.1	10.9	8.4	7.8	0.6	52.2	48.1	4.2
str_m1_o1_o2_f0_pcu.sym.1.out	UO	0.56	3615	2011	0.83	1.49	13.1	10.8	8.5	7.9	0.6	52.2	48.4	3.8
str_m1_o1_o2_f0_pcu.sym.101.out	UO	0.56	3549	1978	0.84	1.50	12.9	10.7	8.5	7.7	0.8	52.1	48.1	4.0
str_m1_o1_o2_f0_pcu.sym.11.out	UO	0.44	4487	1986	0.84	1.89	12.4	10.3	10.4	9.7	0.8	52.0	48.2	3.8
str_m1_o1_o2_f0_pcu.sym.15.out	UO	0.41	4983	2054	0.84	2.04	12.7	9.1	11.1	10.3	0.8	52.0	48.1	3.9
str_m1_o1_o2_f0_pcu.sym.2.out	UO	0.47	4179	1977	0.83	1.75	11.9	9.8	9.8	9.0	0.9	52.0	48.0	3.9
MOF-5									**7.8**			**51.9**		

GCMC and ML generally agree to within 1 wt.% and 3-5 g/L

MOF Synthesis

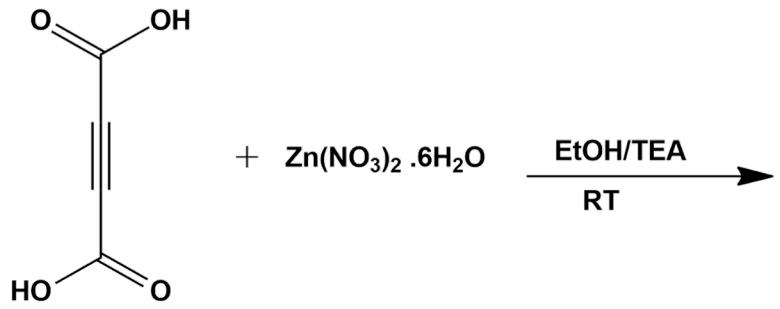
Top performing MOFs predicted by GCM calculations for temperature + pressure swing

Name	Database	Density	GSA	VSA	VF	PV	LCD	PLD	TPS UG	TPS UV
MOF-5_cooh_2_16_4_basic_opt	Mail-order	0.70	3072	2154	0.68	0.68	7.8	12.2	8.0	61.1
MOF-5_cooh_2_2738_1_basic_opt	Mail-order	0.47	4548	2149	0.78	1.34	7.8	15.8	10.8	57.7
BOQQAB (MOF-650)	CSD	0.49	3908	1919	0.85	1.73	18.3	9.9	10.2	56.5
MOF-5_cooh_2_972_1_basic_opt	Mail-order	0.67	3038	2037	0.74	0.95	6.7	11.9	7.5	54.9
hypotheticalMOF_5056615_i_1_j_29_k_2_m_2_cat_1	Northwestern	0.56	4388	2474	0.79	1.41	7.9	9.6	8.6	53.8
ODIXEG (PCN-516)	CSD	0.55	4090	2259	0.84	1.42	10.4	7.5	8.8	53.7
hypotheticalMOF_5057692_i_1_j_29_k_19_m_2	Northwestern	0.55	4546	2489	0.80	1.47	7.2	9.4	8.8	53.6
ENITAX	CSD	0.57	4021	2304	0.83	1.36	10.1	7.2	8.5	53.5
FINJAO	CSD	0.47	6977	3258	0.80	1.70	7.4	6.4	10.2	53.5
TEQPEM	CSD	0.57	3456	1980	0.86	1.45	17.2	9.2	8.5	53.5

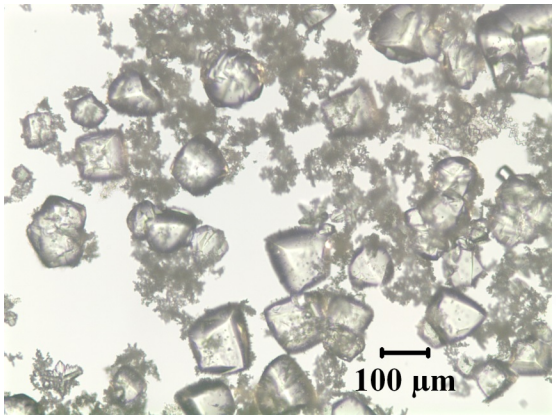
Red highlighted MOFs were chosen for synthesis and activation

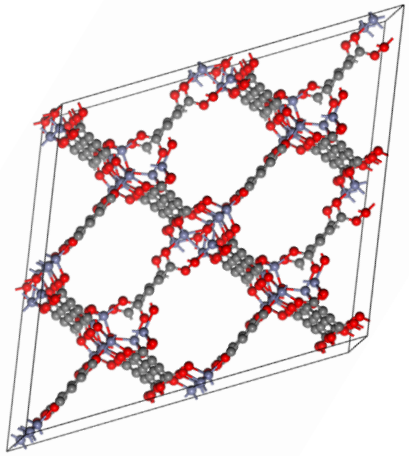
MOF-31 Synthesis

MOF-5_cooh_2_16_4_basic_opt

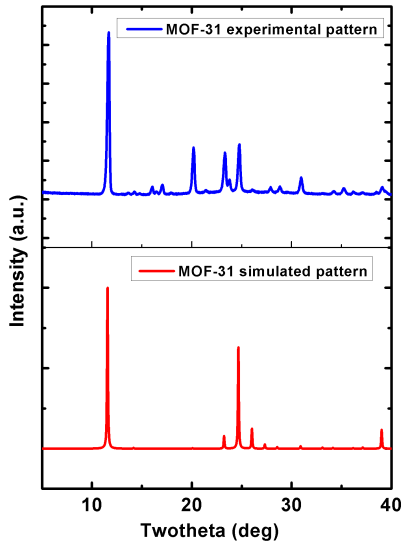


H_2ADC (Acetylene dicarboxylic acid)





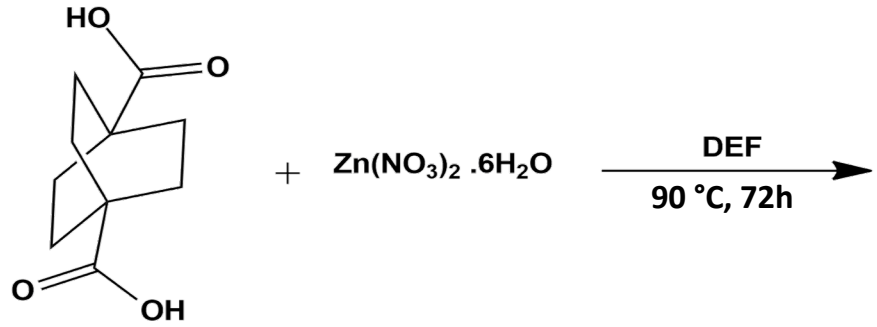
- Obtained material (in combination with a minor phase) washed with ethanol two times and solvent exchange with CH_2Cl_2. Activated by dynamic vacuum (10^{-2} Torr) at room temperature for 24 h.
- The obtained material exhibits a very low BET surface area: 13 m^2/g (calculated value: 3072 m^2/g).
- **Flowing supercritical CO_2 activation:** The ethanol soaked sample was exchanged with supercritical CO_2 at 100 bar resulting in a material with a very low BET surface area of 13 m^2/g.



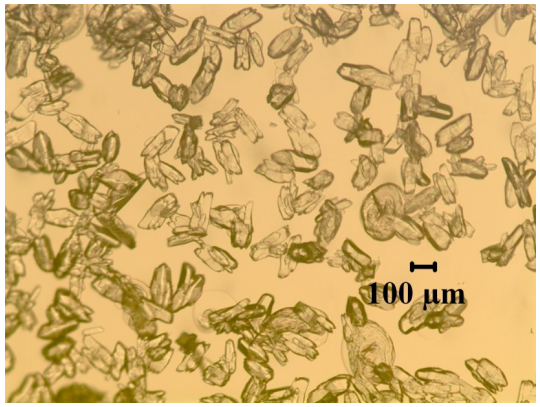
O. Yaghi, et al., J. Am. Chem. Soc., **123**, 8239(2001)

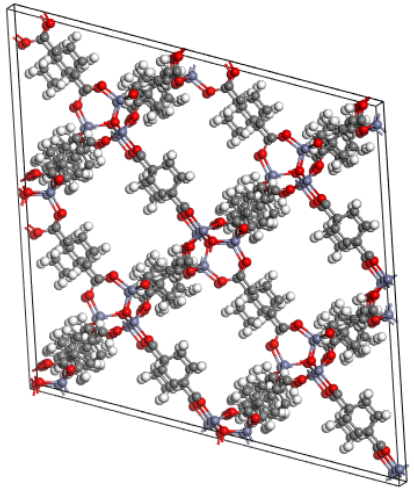
TMOF-2 Synthesis

MOF-5_cooh_2_972_1_basic_opt

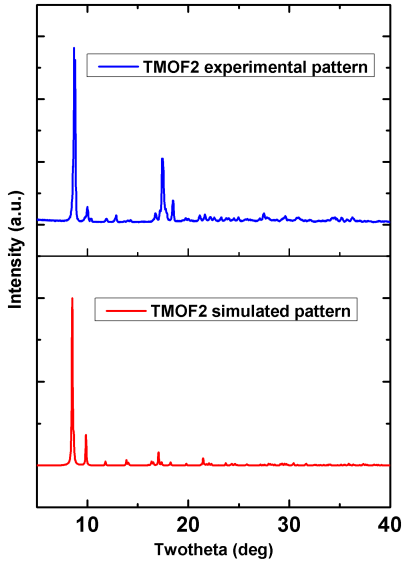


H_2BCCA (Bicyclo[2.2.2]octane-1-carboxylic acid)





- Desired MOF solvent exchanged with DMF followed by CH_2Cl_2. Activated by dynamic vacuum (10^{-2} Torr) at room temperature for 24 h.
- The obtained material exhibits a very low BET surface area of 8 m^2/g (calculated value: 3038 m^2/g).
- **Flowing supercritical CO_2 activation:** The DMF soaked sample was exchanged with supercritical CO_2 at 100 bar resulting in a material with a BET surface area of 621 m^2/g.



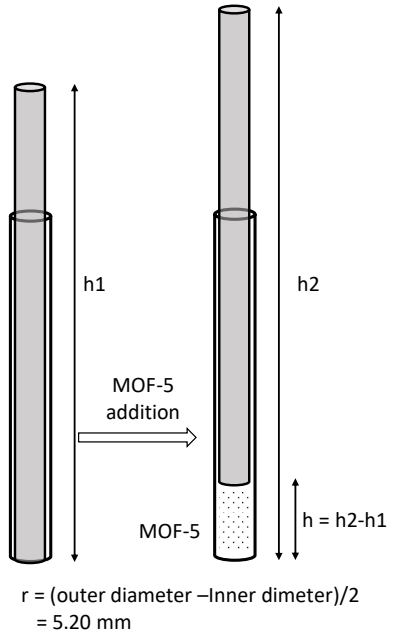
PJ Llabres-Campaner, et al., Dalton Trans., **46**, 7397(2017)

Milestone 2.2

Task No.	Task Title	Milestone Description	Milestone Verification Process	Due (Month)	Status
2.2	Particle size control	Determine if particle size influences packing efficiency by more than 10% for particles whose size varies by more than an order of magnitude	Void fraction measurements	15	**Complete**

Packing Density Measurements

- Preliminary MOF-5 packing density measurements were performed using the setup below in a glovebox
- The packing efficiencies were assessed experimentally and expressed as a fraction of the single crystal density
- Around 100 mg of MOF-5 sample was used for each measurement, which was tapped ~100 – 200 times until no visible gaps were seen. Finally, a plunger was used to determine the upper position of the powder column



Particle Size (MOF-5)	Packing Density (g/cm^3)	% SC Density
100-400 µm	0.345	57.2
500-1000 µm	0.372	61.8
1300-1900 µm	0.366	60.7

Fig. 1: Setup used for determining packing density. Volume of sample used was determined using the formula, volume of cylinder = $\pi * r^2 * h$. Efforts were made to be as consistent as possible for each measurement, for instance, the number of taps, the amount of sample used were all kept as constant as possible.

Packing Density Measurements

To achieve consistent results we designed a small-scale jolting volumeter for packing density determination

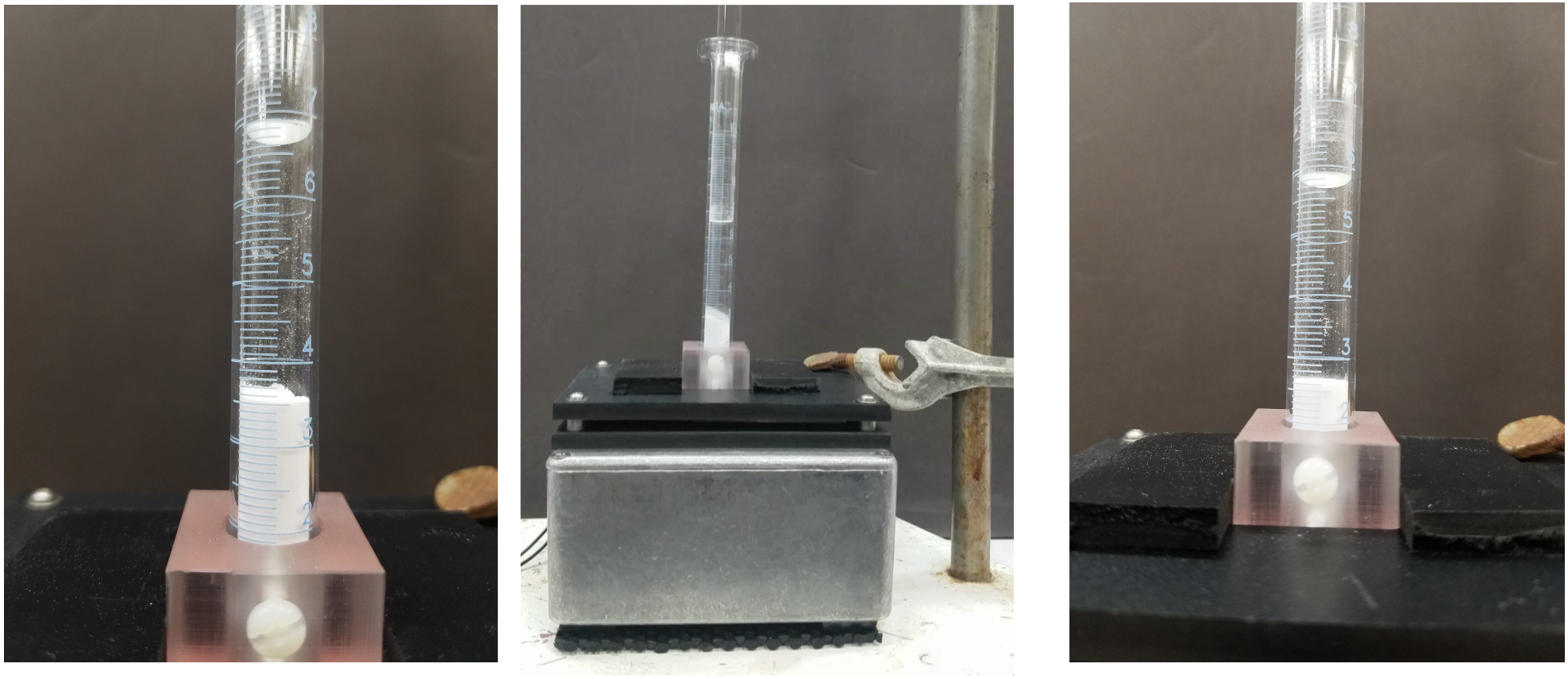
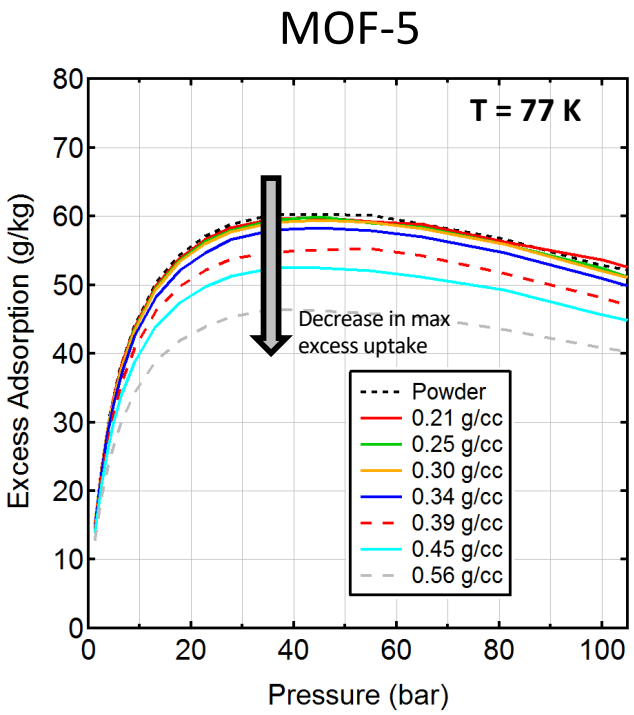


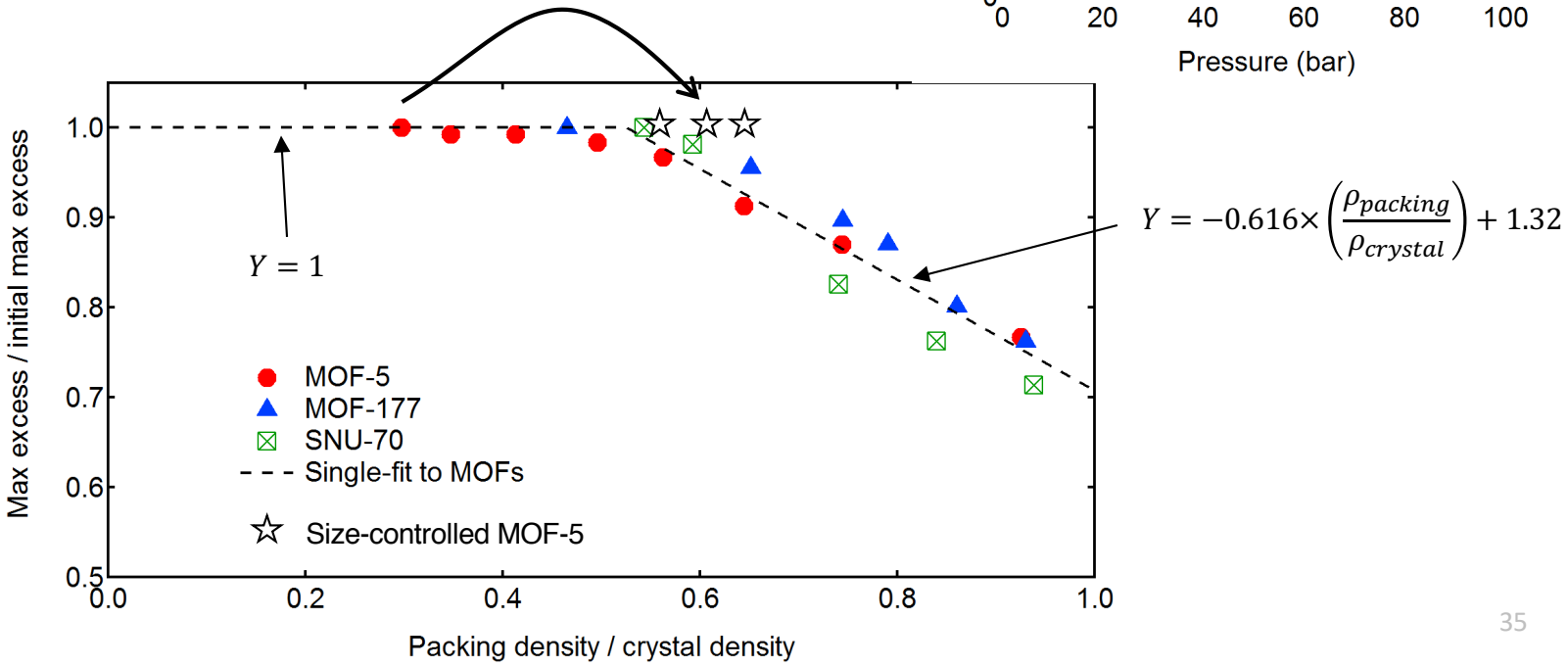
Fig. 1: Instrumental setup for determining packing density for NaCl

Compaction Properties of MOFs

Controlling MOF particle size shows promise for maintaining capacity at higher packing densities, without the need for physical compaction

MOF-5 Particle Size	Packing Density (g/cm³)	% SC Density
100-400 µm	0.345	57.2
500-1000 µm	0.372	61.8
1300-1900 µm	0.366	60.7





Summary

- **Goal**: Overcome volumetric limitations associated with physisorptive hydrogen storage at both the materials and systems level in metal-organic frameworks (MOFs)
- **Approach**:
  - Control MOF crystal morphology and crystallite size distribution to increase packing density
  - Apply machine learning techniques to identify, design, and demonstrate high-capacity MOFs
- **Accomplishments:**
  - Used machine learning to identify MOF crystallographic properties consistent with a target hydrogen storage capacity (1st Go/No-go)
  - Extended ML predictions of MOF capacity to temperature + pressure swing operation
    - Highest-capacity MOFs were validated by more rigorous GCMC calculations
    - Experimental synthesis of selected MOFs is on-going
  - Controlling MOF particle size shows promise for maintaining capacity at higher packing densities, without the need for physical compaction

UNIVERSITY OF MICHIGAN
umich.edu/~djsiege
djsiege@umich.edu

Potential Future Work

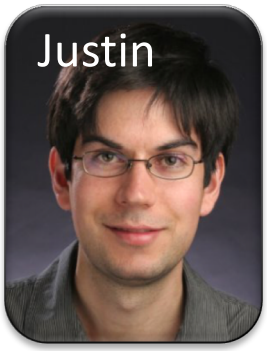
Focus on 2nd go/no-go milestone: Demonstrate an improvement in either

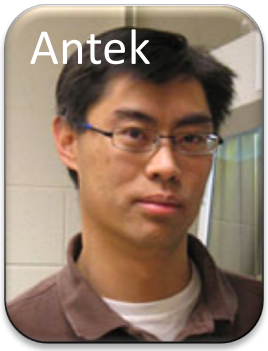
a) a MOF with a single crystal volumetric capacity greater than 39 g/L usable capacity measured at 77 K, and 5-100 bar pressure (i.e., a 10% increase over the current state-of-the-art NU-100) through Machine Learning-directed material development

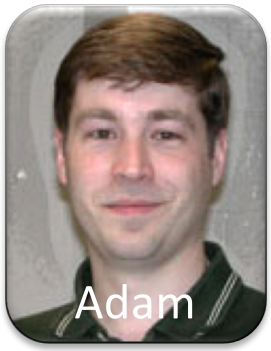
OR

b) a 15% increase in tap density through crystal engineering methods for a specific MOF compared to its non-optimized powder, with a minimal loss in surface area

The Team







Collaborations

University of Michigan, Mechanical Engineering

- Atomistic simulation and project management

University of Michigan, Dept. of Chemistry

- Synthesis and characterization of targeted MOFs

Ford Motor Company (sub-contractor)

- PCT measurements
- Materials augmentation, characterization, scale-up, and system modeling

HSECoE/SRNL (unfunded collaborator)

- Assistance with system models (David Tamburello)

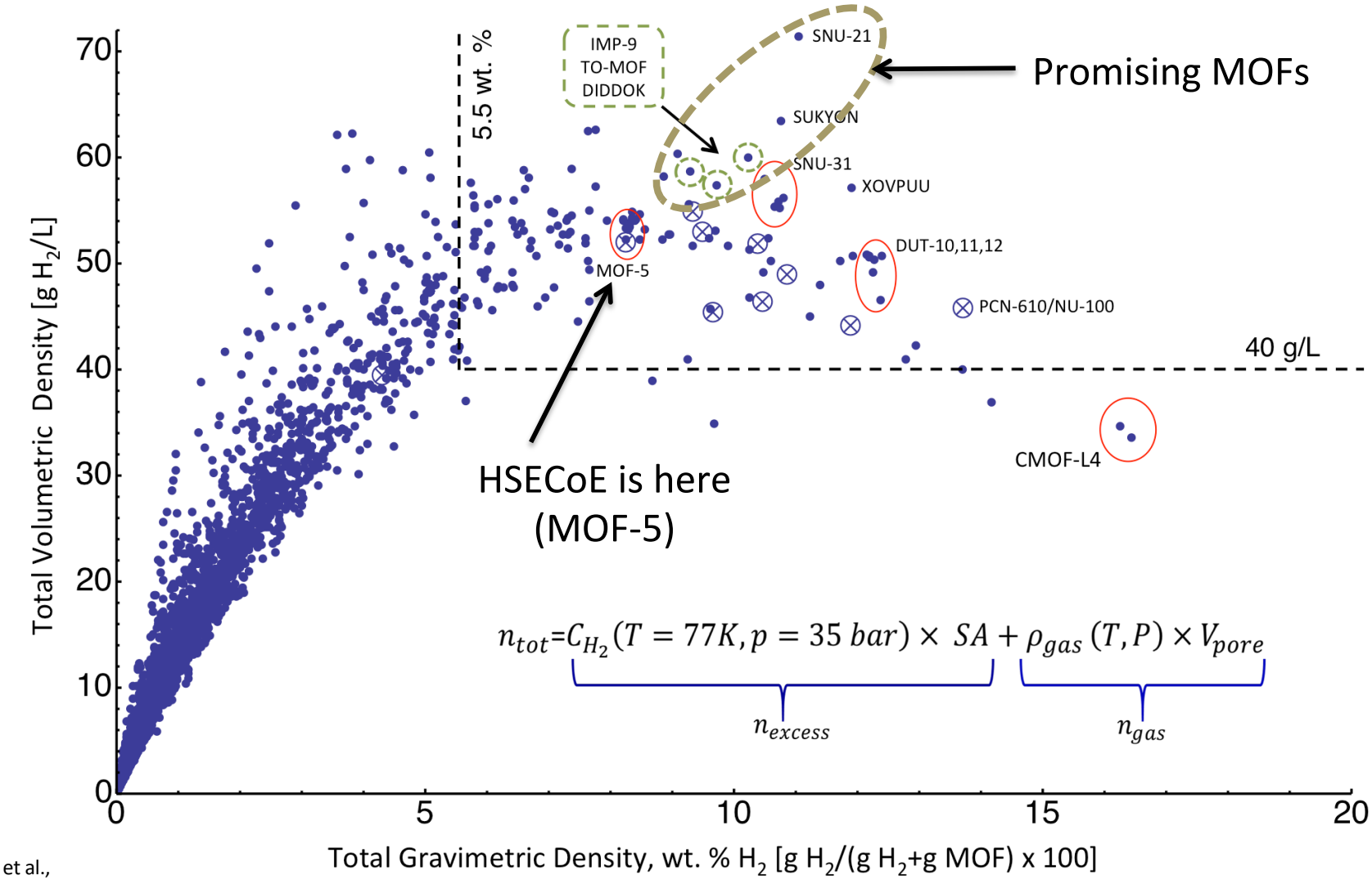
Challenges and Barriers

- Many more compounds identified by computation than can be synthesized
  - Assessment by a human is needed before synthesis can proceed
  - This is a bottleneck
- Structure collapse or incomplete solvent removal during activation
  - “Can it be made?”
  - Failure to achieve expected surface area and porosity
  - Properties that control “synthesizability” are not well-understood
- Incorrect, incomplete, or disordered crystal structure data
  - Garbage in, garbage out
  - False positives in screening

Technical Backup Slides

High-throughput Screening

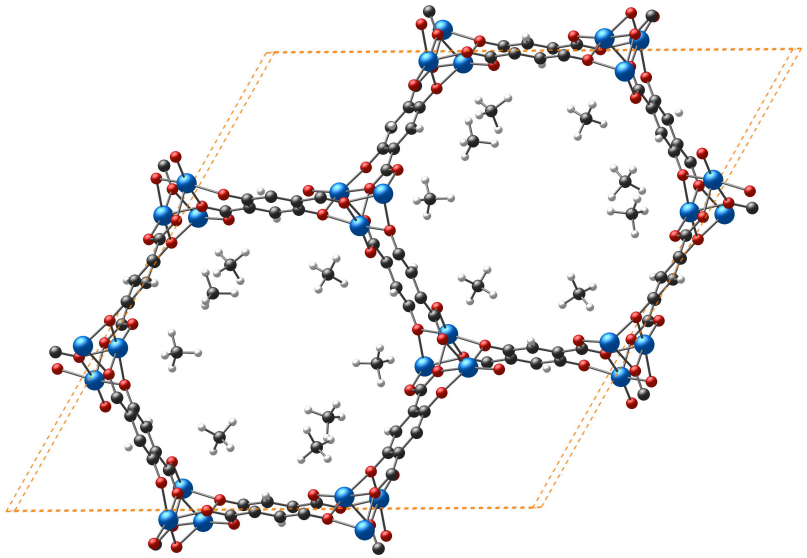
Prior work: developed a database of MOFs by mining the CSD. Chahine rule and crystal structure were used to predict H_2 capacity in thousands of compounds



J. Goldsmith, et al.,
Chem. Mater., **25**, 3373 (2013).

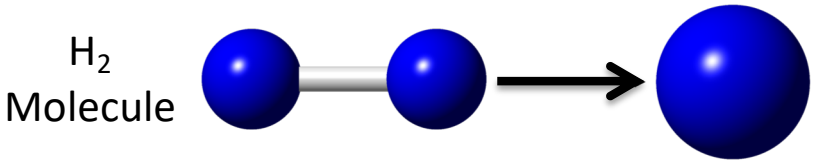
Grand Canonical Monte Carlo

- GCMC = atomistic method that calculates the total amount of H_2 (adsorbed + gas phase) contained within the pore space of a MOF at given T, P
- Does not rely on empirical correlations such as the Chahine-rule



Example GCMC simulation of CH_4 adsorption in Ni-DOBDC at 298 K and 35 bar

- Calculations employ the MGS* and the **Pseudo-FH**** unified atom models for H_2-MOF interactions
- MOF atoms are fixed



$$U_{ij}(r_{ij}) = 4\varepsilon\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

Force Field	Sigma (Å)	Epsilon/k_B (K)
MGS	2.958	36.7
Pseudo-FH	3.064	30.1

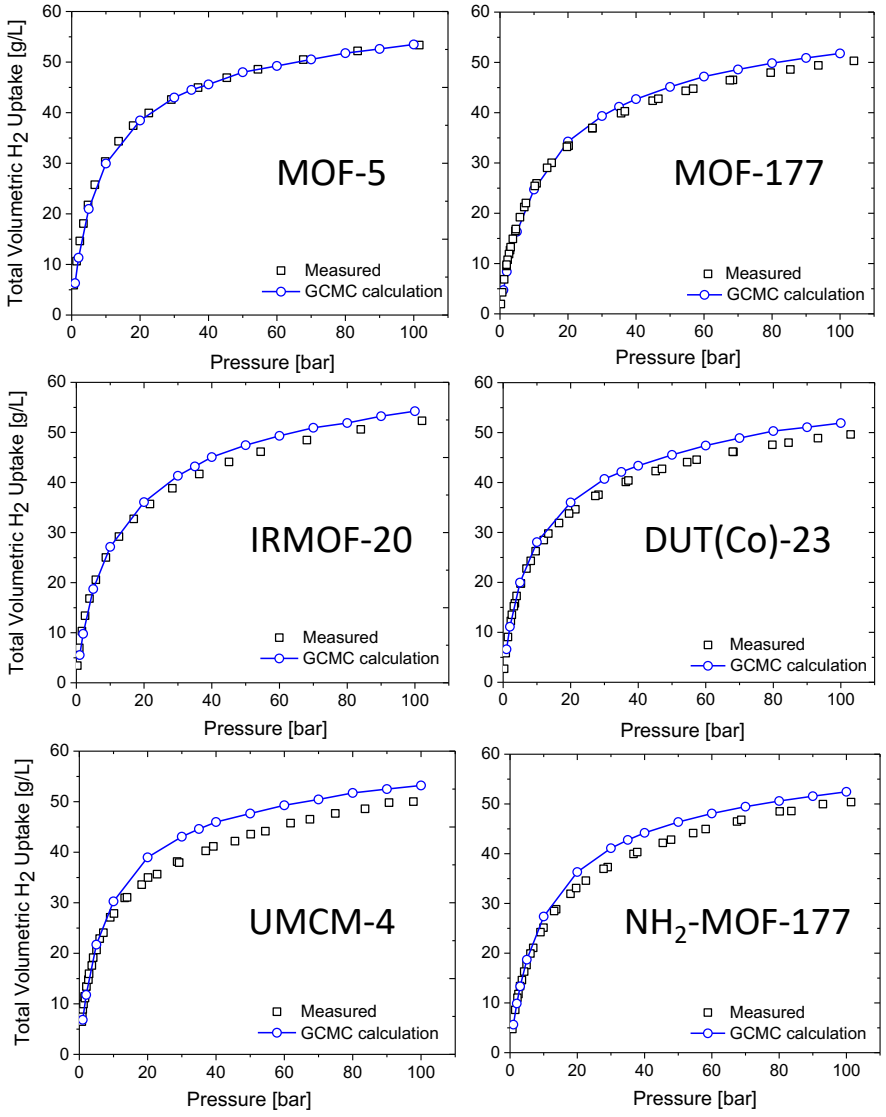
*Michels, de Graaff and Seldam, *Physica*, **1960**, 26, 393; Ryan, Broadbelt, and Snurr, *Chem. Comm*. 2008, 4134

Fischer, Hoffmann, Fröba, *ChemPhysChem*, **2009,10, 2647.

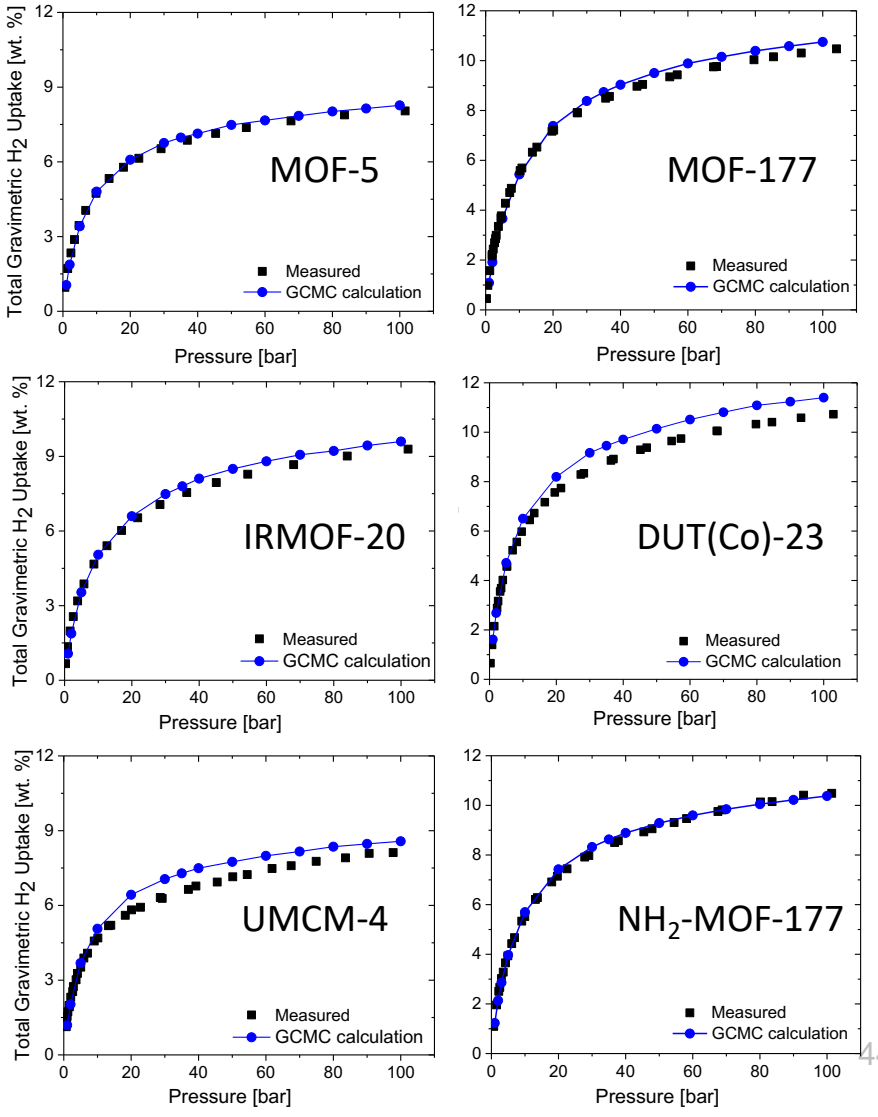
Examples of Simulated Isotherms

GCMC isotherms calculated with the pseudo-Feynman-Hibbs interatomic potential are in very good agreement with our measured isotherms

Total Volumetric H_2 Uptake

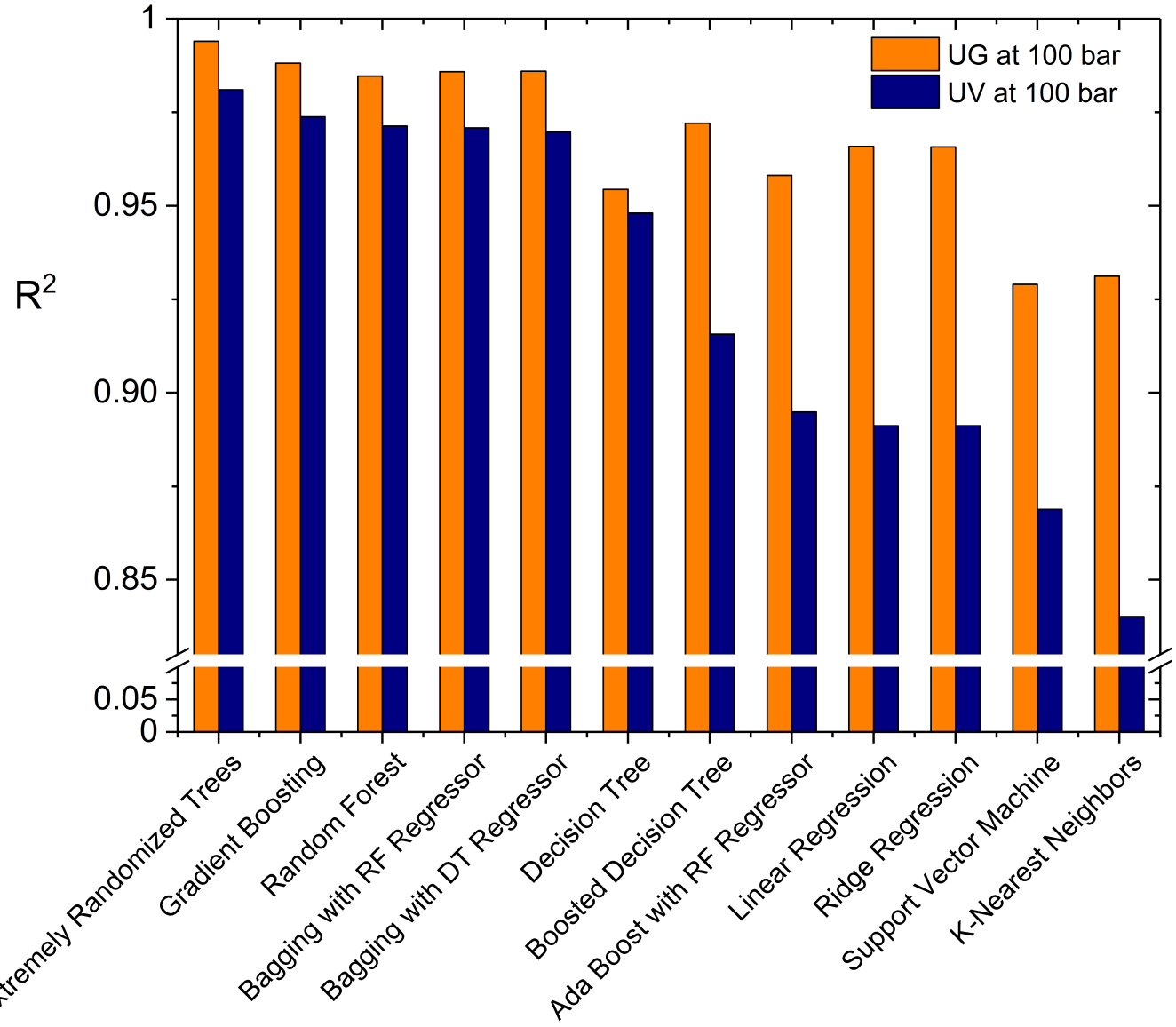


Total Gravimetric H_2 Uptake



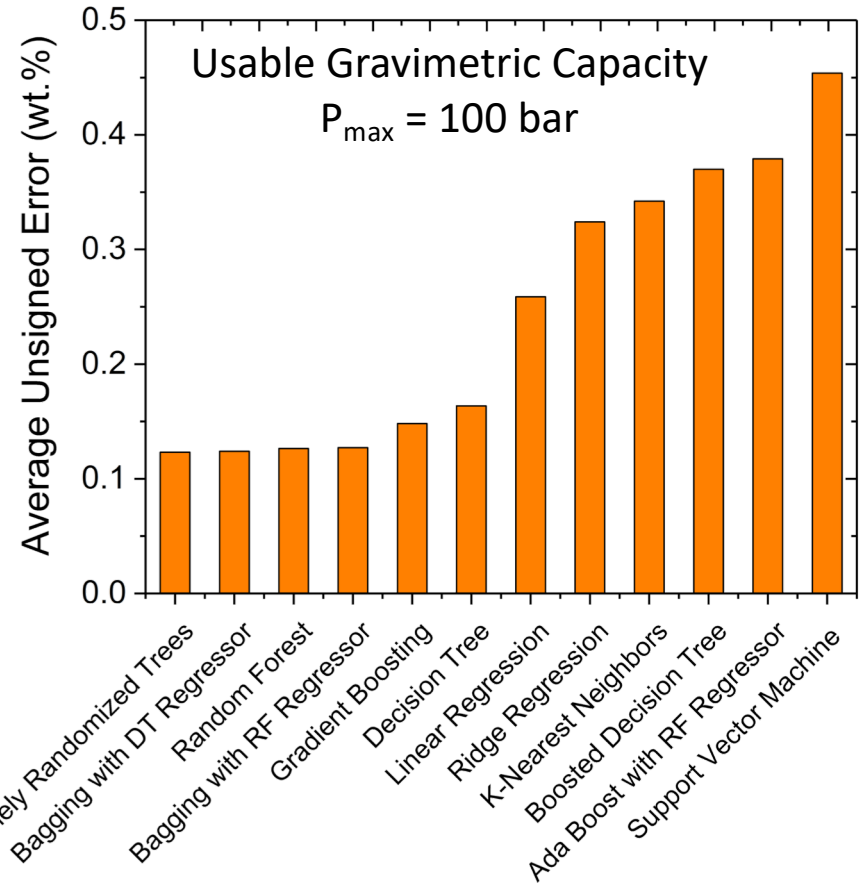
Benchmarking ML Methods

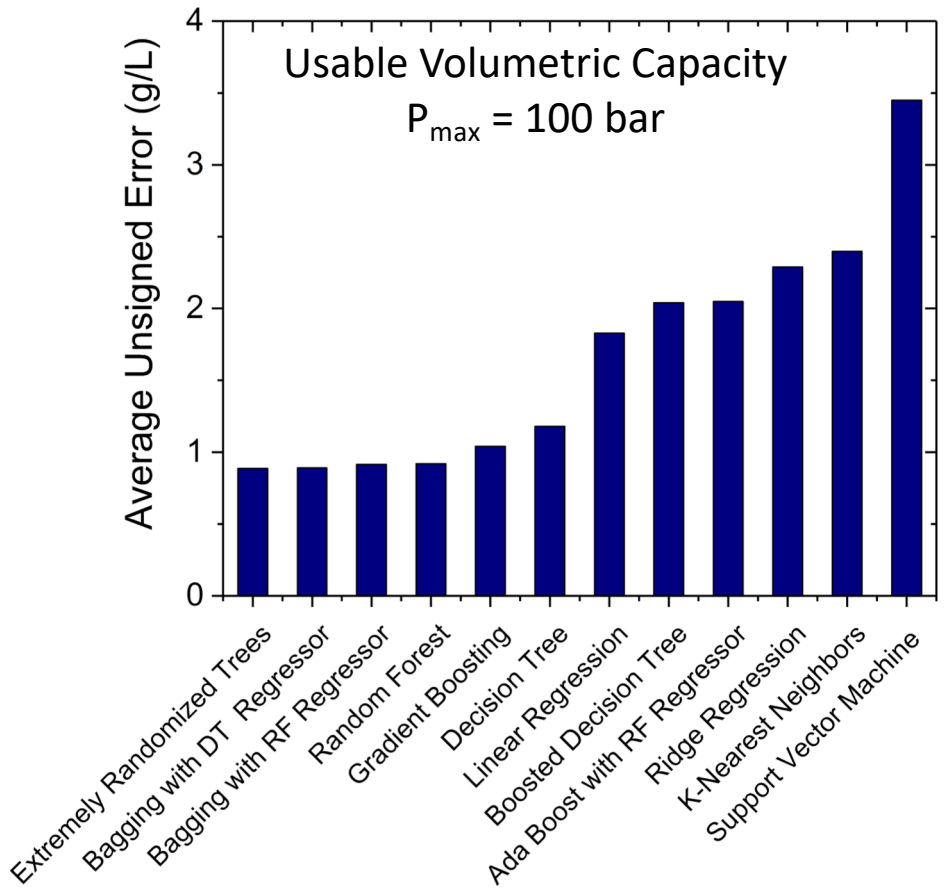
The Extremely Randomized Trees method is the best performing ML algorithm



Comparison of ML Methods

The Extremely Randomized Trees method is the best performing ML algorithm

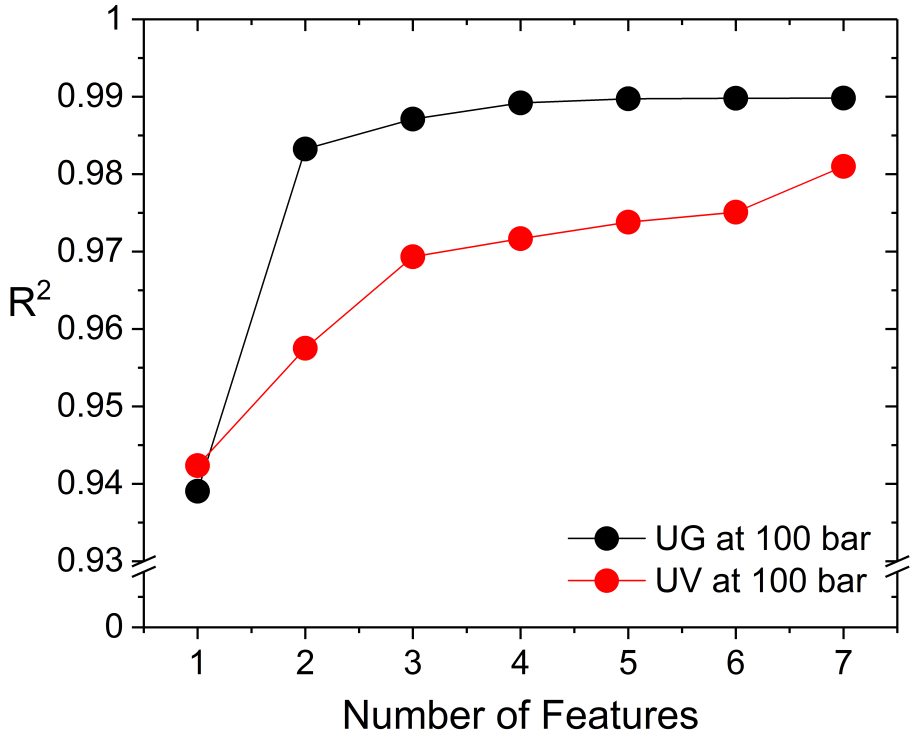




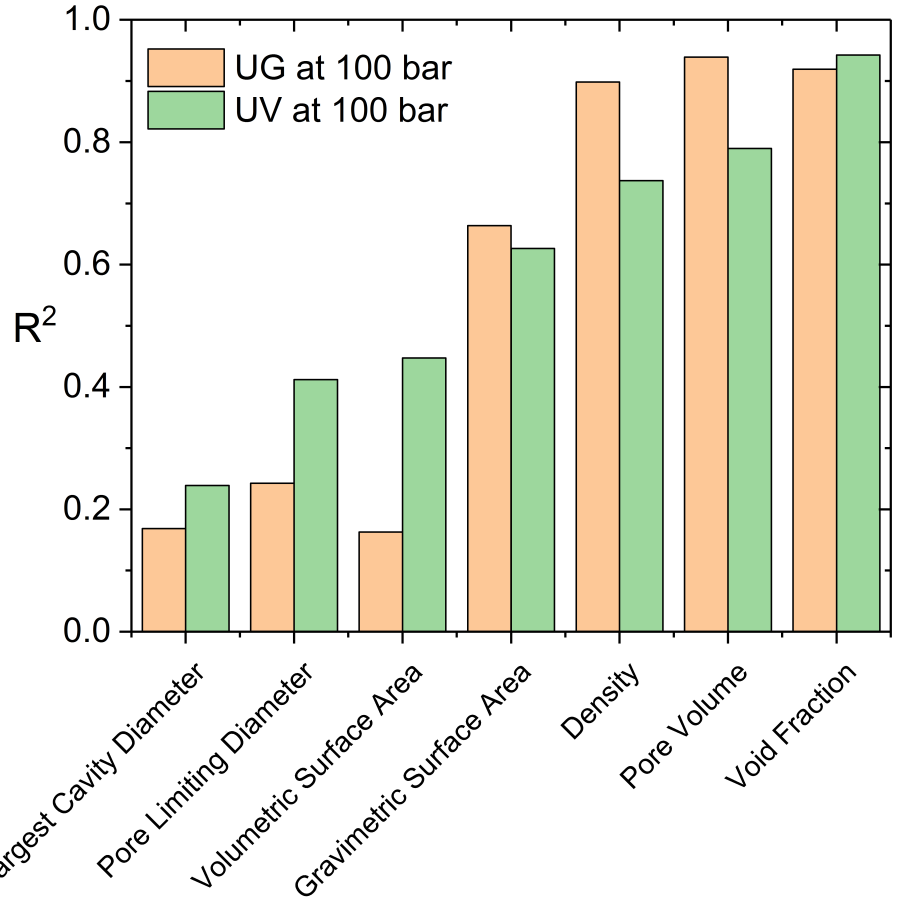
Structure-Property Correlations

- Void fraction shows the strongest correlation with UV
- Only 4 features needed to predict UV with over 96% accuracy

Effect of Number of Features

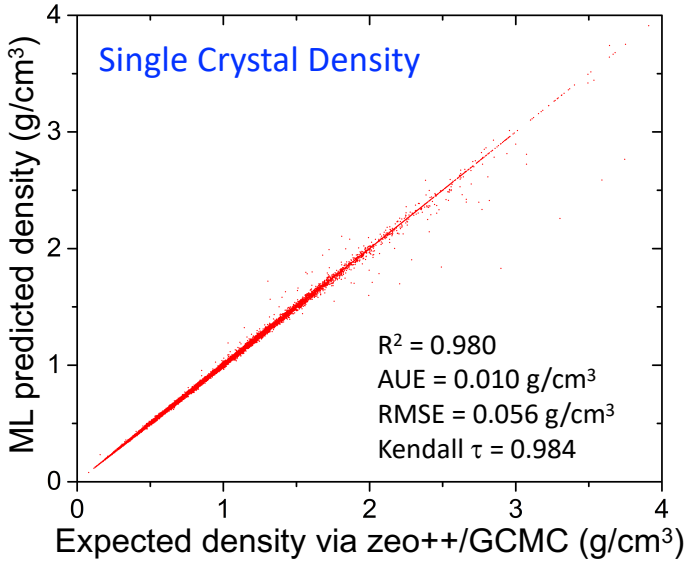


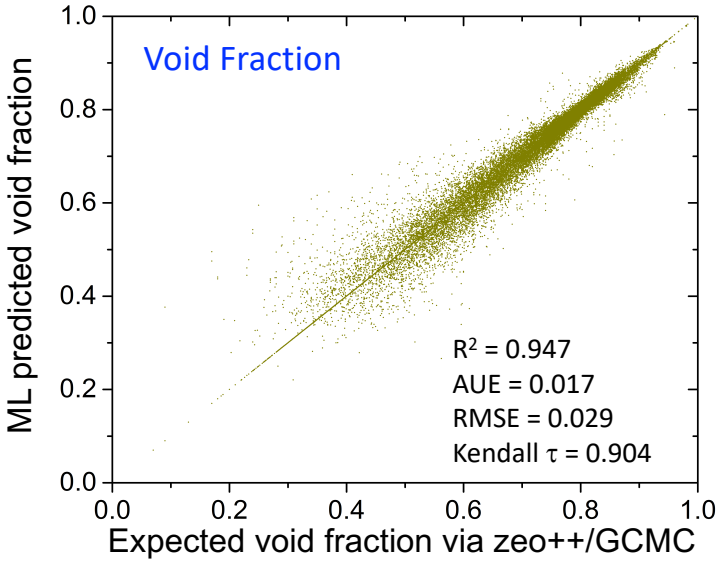
Single Feature Performance

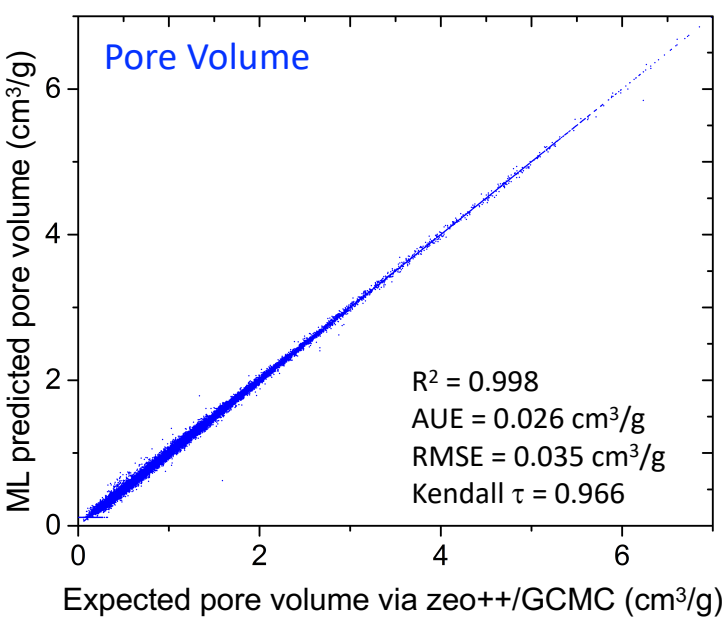


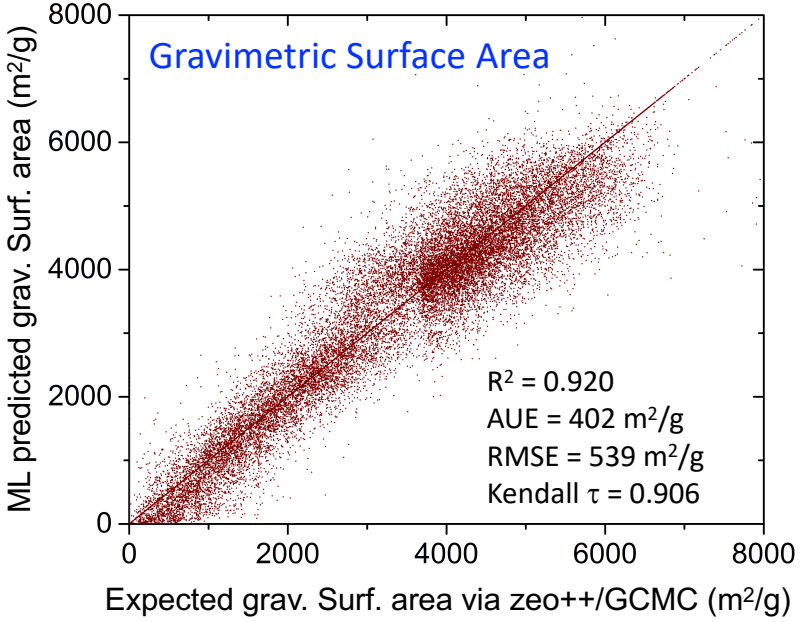
Each point on the plot represents the highest R^2 value among all possible(2^n -1) combinations of (n = 1,2,3,4,5,6,7) features

Go/No-Go


Single Crystal Density
ML predicted density (g/cm³)
Expected density via zeo++/GCMC (g/cm³)
R² = 0.980
AUE = 0.010 g/cm³
RMSE = 0.056 g/cm³
Kendall τ = 0.984

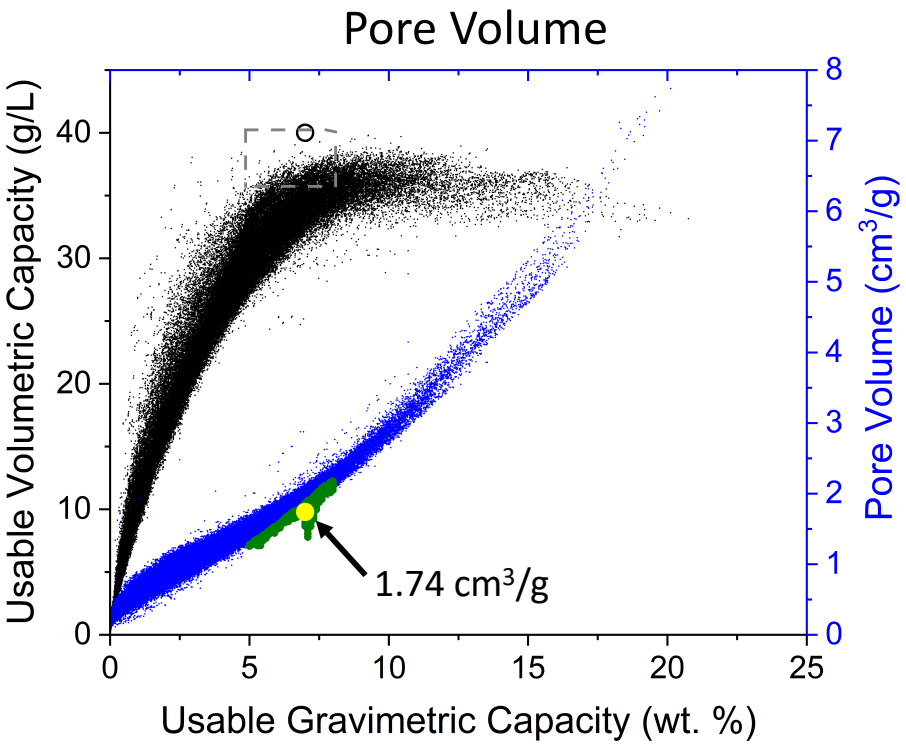

Void Fraction
ML predicted void fraction
Expected void fraction via zeo++/GCMC
R² = 0.947
AUE = 0.017
RMSE = 0.029
Kendall τ = 0.904

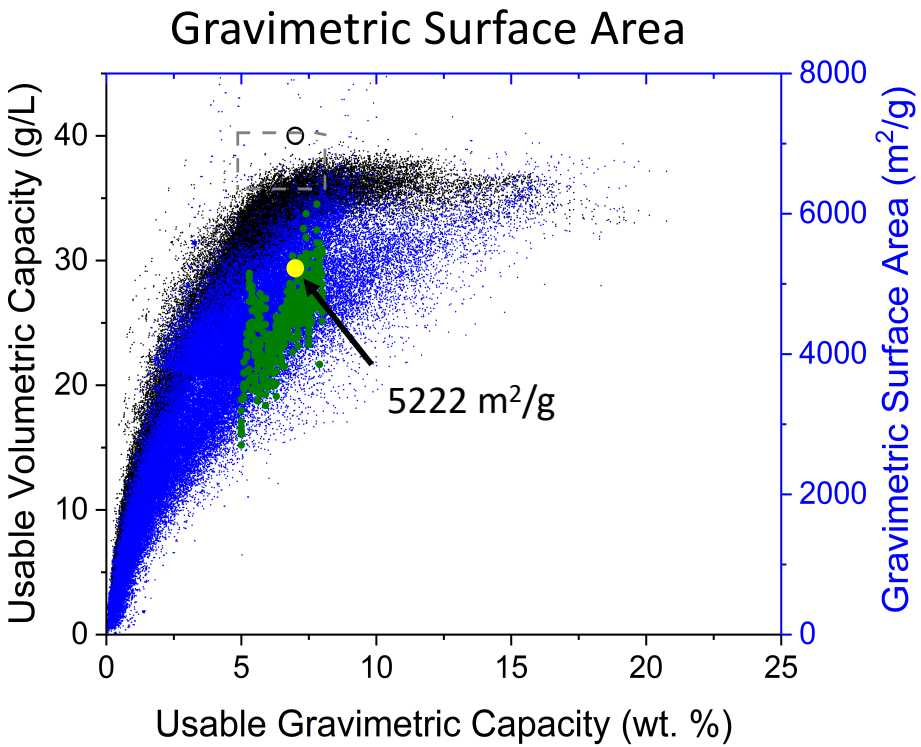

Pore Volume
ML predicted pore volume (cm³/g)
Expected pore volume via zeo++/GCMC (cm³/g)
R² = 0.998
AUE = 0.026 cm³/g
RMSE = 0.035 cm³/g
Kendall τ = 0.966

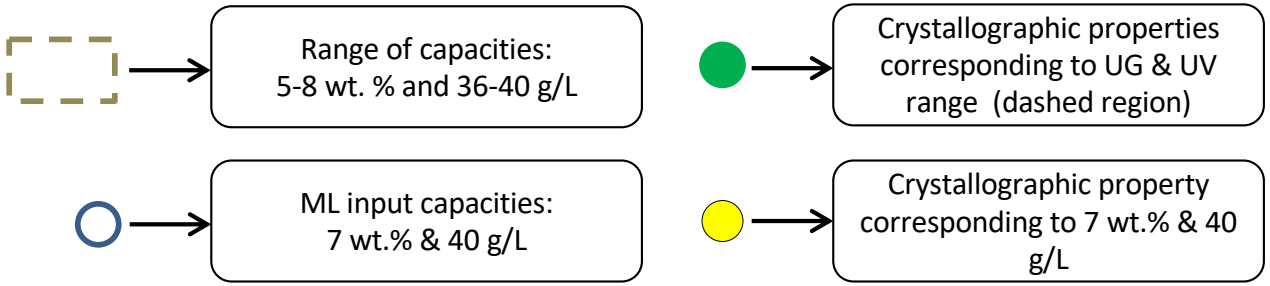

Gravimetric Surface Area
ML predicted grav. Surf. area (m²/g)
Expected grav. Surf. area via zeo++/GCMC (m²/g)
R² = 0.920
AUE = 402 m²/g
RMSE = 539 m²/g
Kendall τ = 0.906

Go/No-Go

ML models were trained 'in reverse' to predict the crystallographic propoperties that correspond to a specified usable capacity





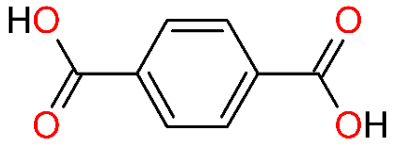


Crystal Size Control (1)

Successfully developed synthesis protocols that produce MOF-5 with different average sizes

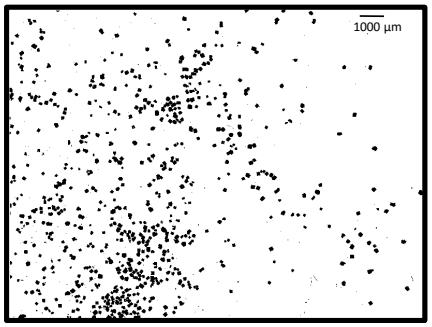
Synthesis Protocols

Varying Metal:Linker ratio

$Zn(NO_3)_2 \cdot 6H_2O$ + 

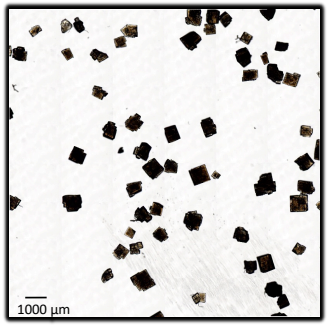
DEF, 100 °C (12hr)

M:L=1.5



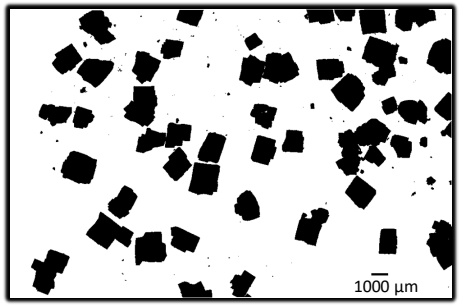
100-400 µm

M:L=2.5



500-1000 µm

M:L=3.5



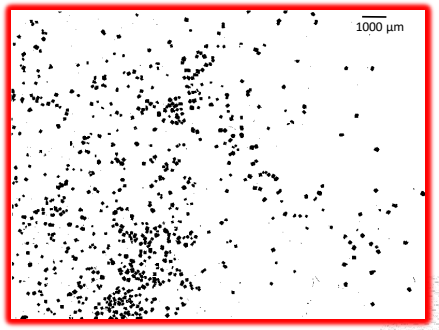
1300-1900 µm

Agitating reaction mixture — M:L=1.5 — 60 rpm → **100-250 µm**; 100 rpm → **50-150 µm**

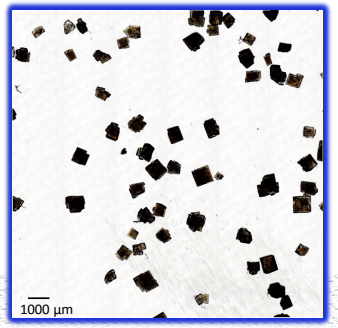
Crystal Size Control (2)

Crystal size histograms indicate the successful control over crystal size

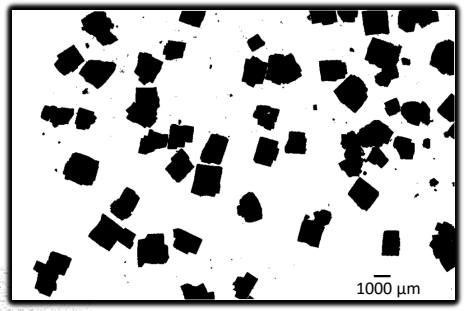
Varying the size distribution of cubic MOF-5 crystallites



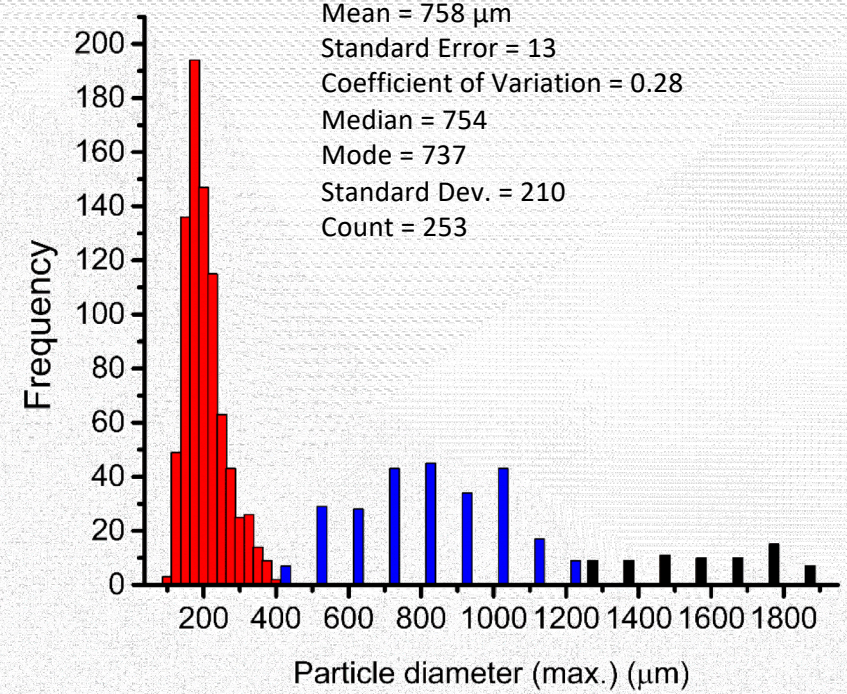
Mean = 192 µm
Standard Error = 1.9
Coefficient of Variation = 0.29
Median = 179
Mode = 174
Standard Dev. = 56
Count = 828



Mean = 758 µm
Standard Error = 13
Coefficient of Variation = 0.28
Median = 754
Mode = 737
Standard Dev. = 210
Count = 253



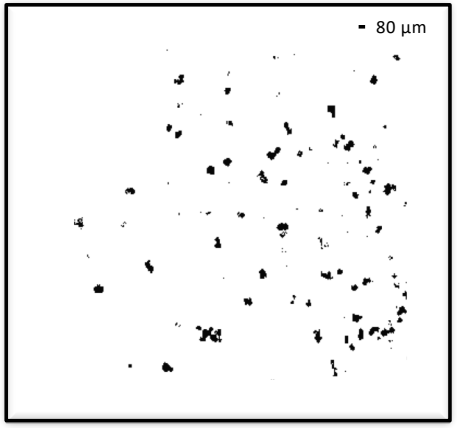
Mean = 1633 µm
Standard Error = 39
Coefficient of Variation = 0.24
Median = 1590
Mode = 1675
Standard Dev. = 389
Count = 97

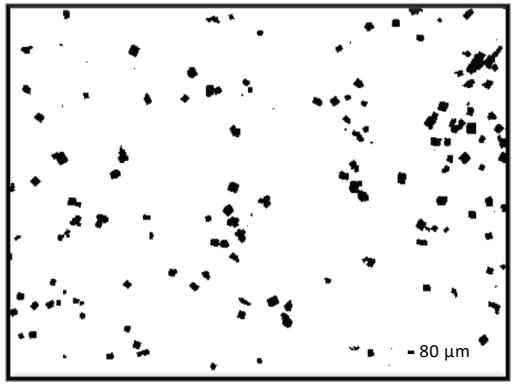


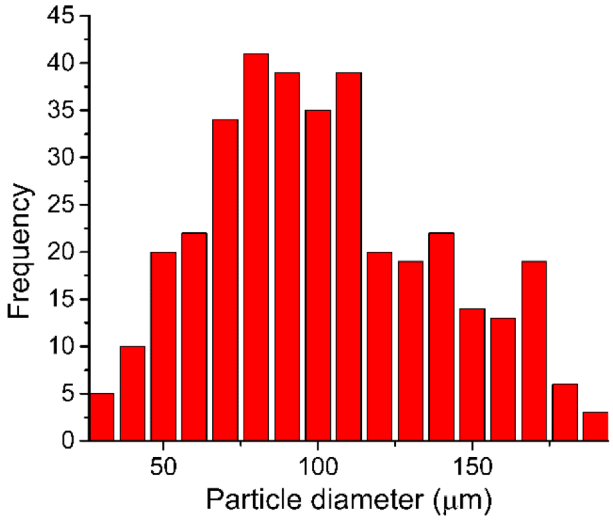
Crystal Size Control (3)

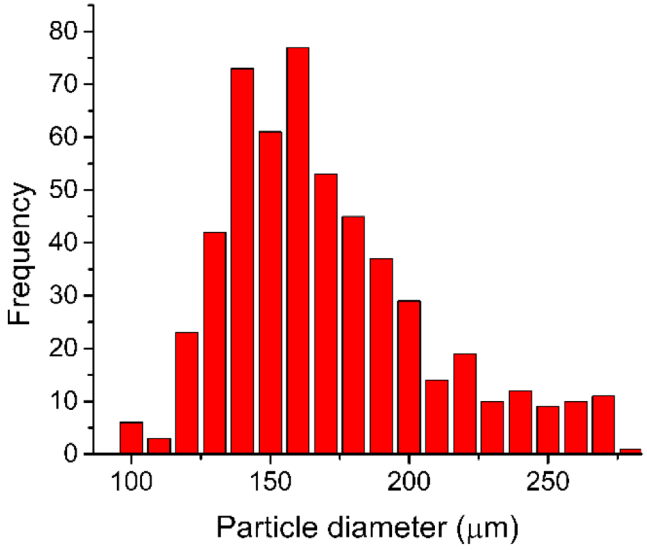
Crystal size histograms indicate the successful control over crystal size

Varying the size distribution of cubic MOF-5 crystallites



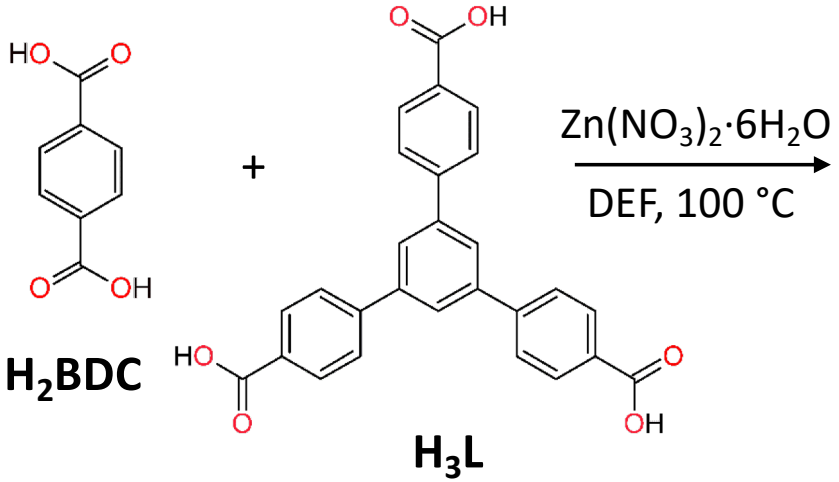






Morphology Control (1)

Identified additive capable of controlling morphology of MOF-5 crystals



H_3L=1,3,5-tris(4-carboxyphenyl)benzene

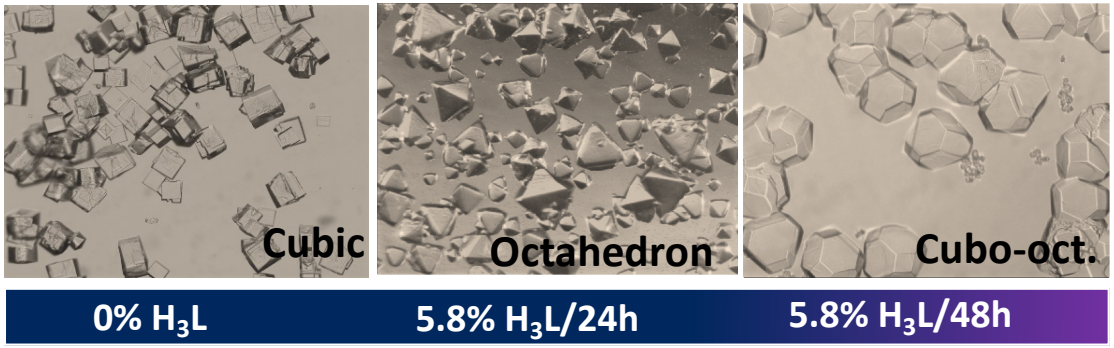


Fig. 1: Optical images of different morphologies of MOF-5 crystals obtained by the addition of H_3L

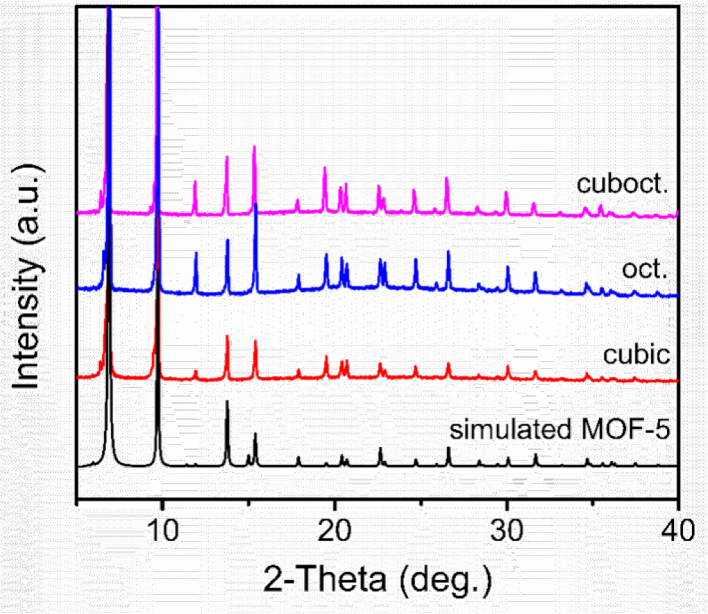
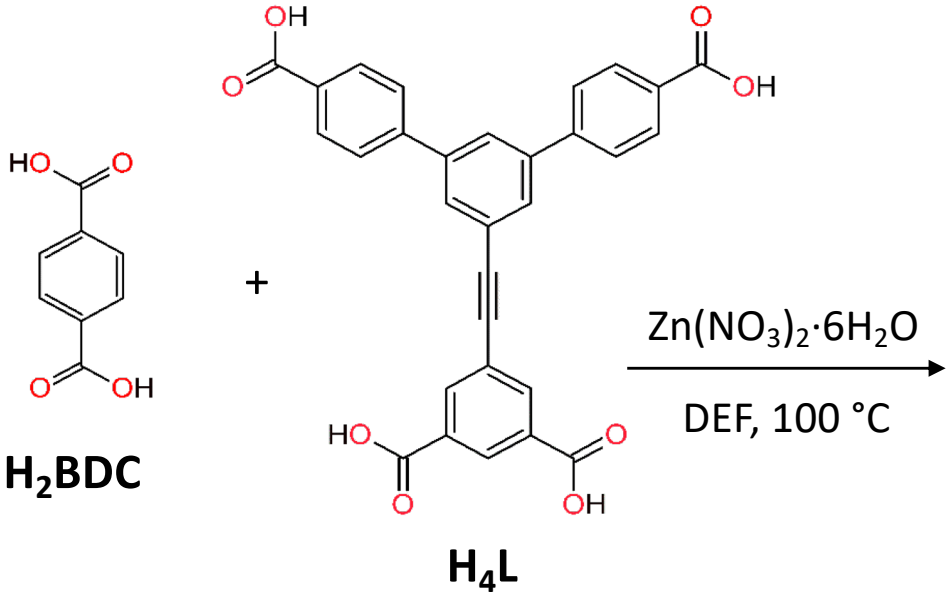


Fig.2: Phase purity of obtained morphologies was confirmed through PXRD and it was observed that powder patterns of all morphologies were found to agree with the simulated pattern of MOF-5 extracted from the crystal structure

Morphology Control (2)

Identified additive capable of controlling morphology of MOF-5 crystals



H_2BDC

H_4L

H_4L=5’-((3,5-dicarboxyphenyl)ethynyl)-[1,1’:3’,1’’-terphenyl]-4,4’’-dicarboxylic acid)

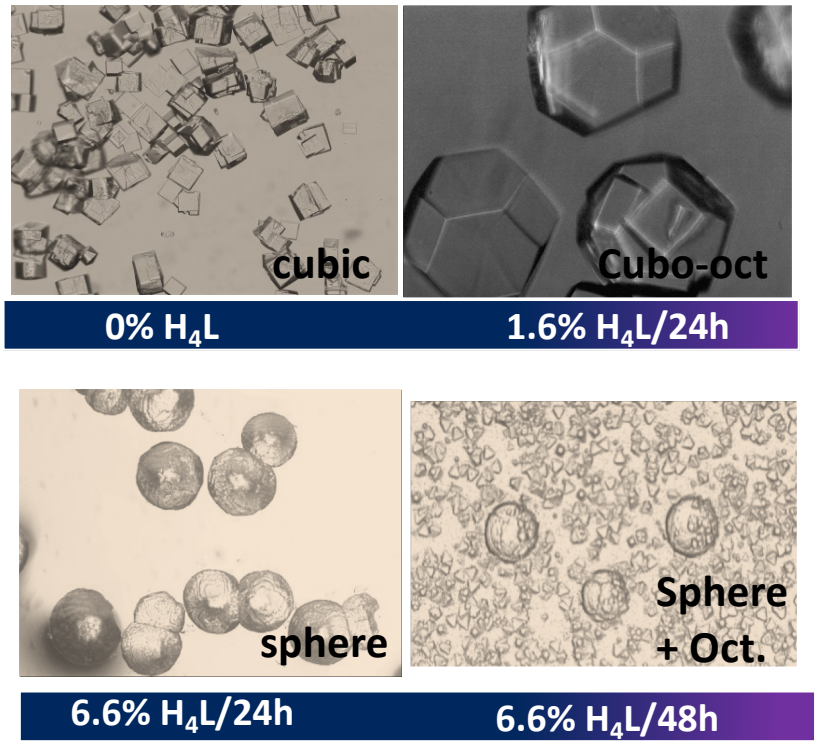


Fig.1: Optical images of different morphologies of MOF-5 crystals obtained by the addition of H_4L

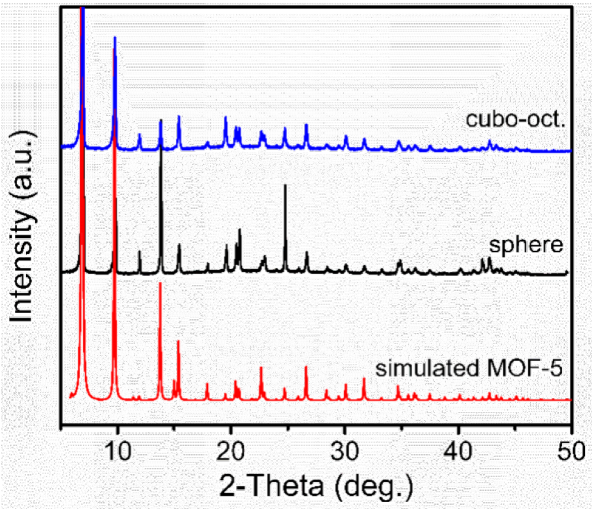
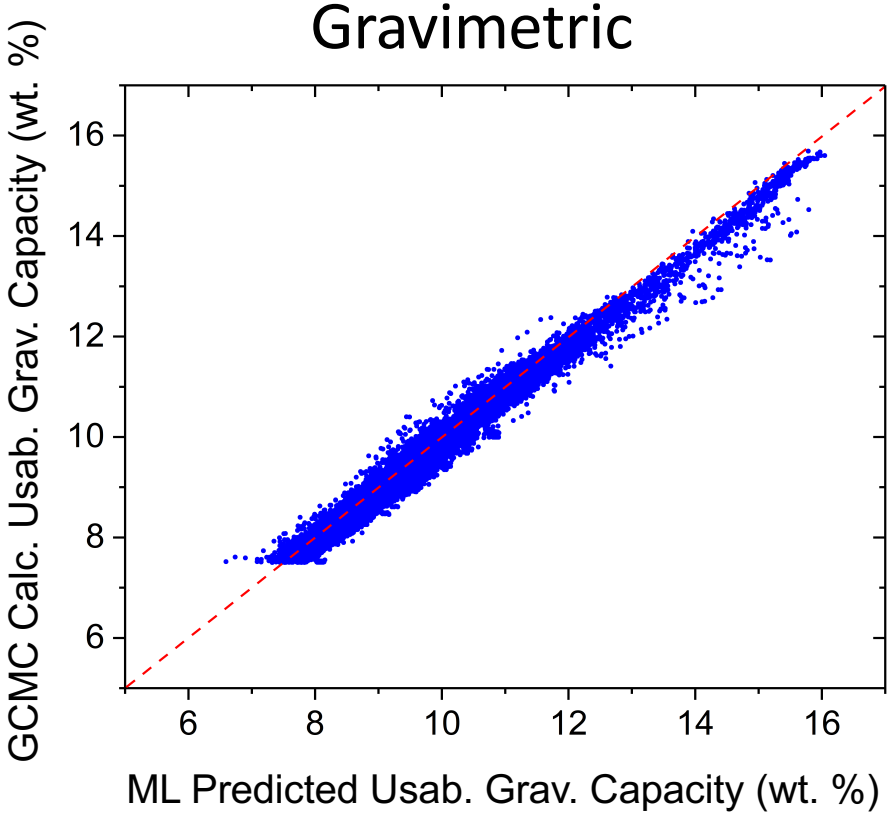


Fig.2: Phase purity of obtained morphologies was confirmed through PXRD and it was observed that powder patterns of all morphologies were found to agree with the simulated pattern of MOF-5 extracted from the crystal structure

PS: ML Predictions and Validations (2)

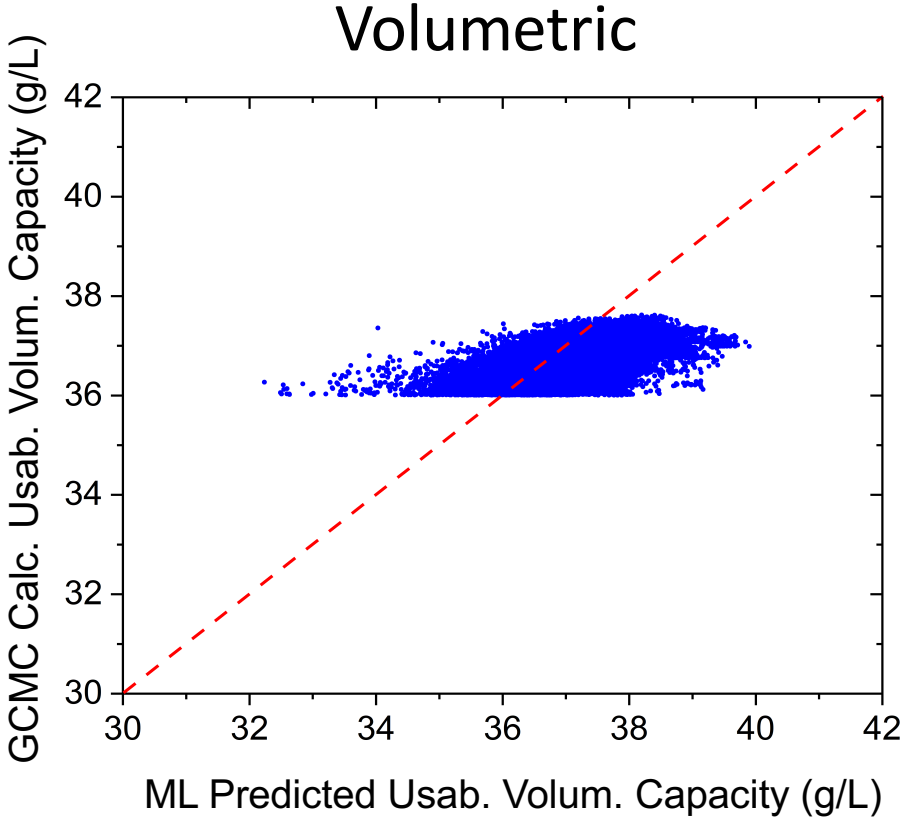
Correlations between ML predicted and GCMC calculated usable capacities of top 15,902 unseen (out of ~400,000 compounds not used in training and testing of ML models) high-capacity MOFs at 77K for the pressure swing between 100 and 5bar

Gravimetric



Maximum Difference	1.67 wt. %
Minimum Difference	-0.96 wt.%
Average Unsigned Error	0.21 wt. %
Standard Deviation	0.17 wt. %

Volumetric

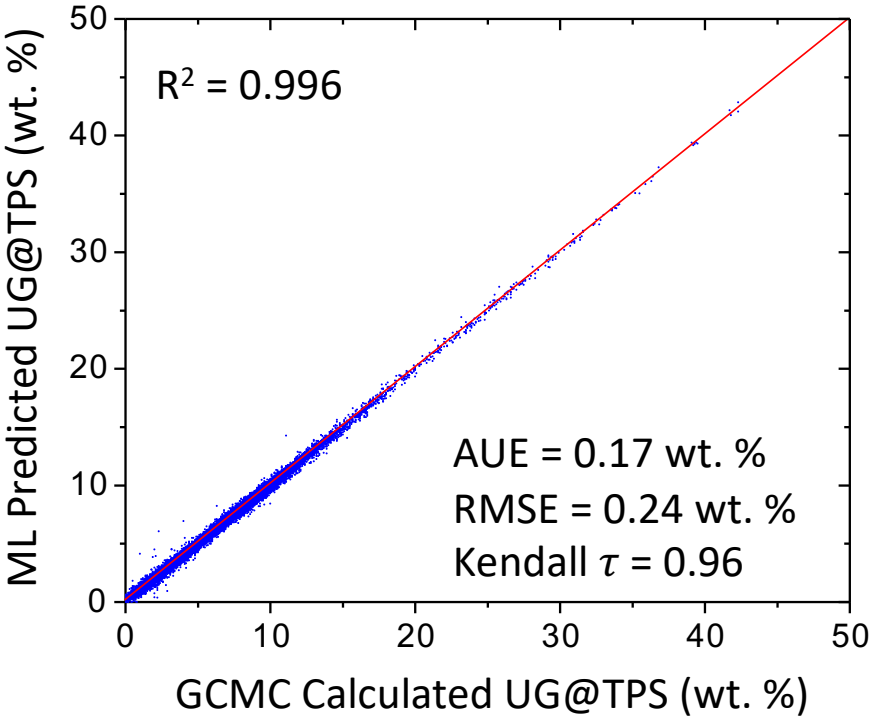


Maximum Difference	3.05 g/L
Minimum Difference	-4.46 g/L
Average Unsigned Error	0.62 g/L
Standard Deviation	0.51 g/L

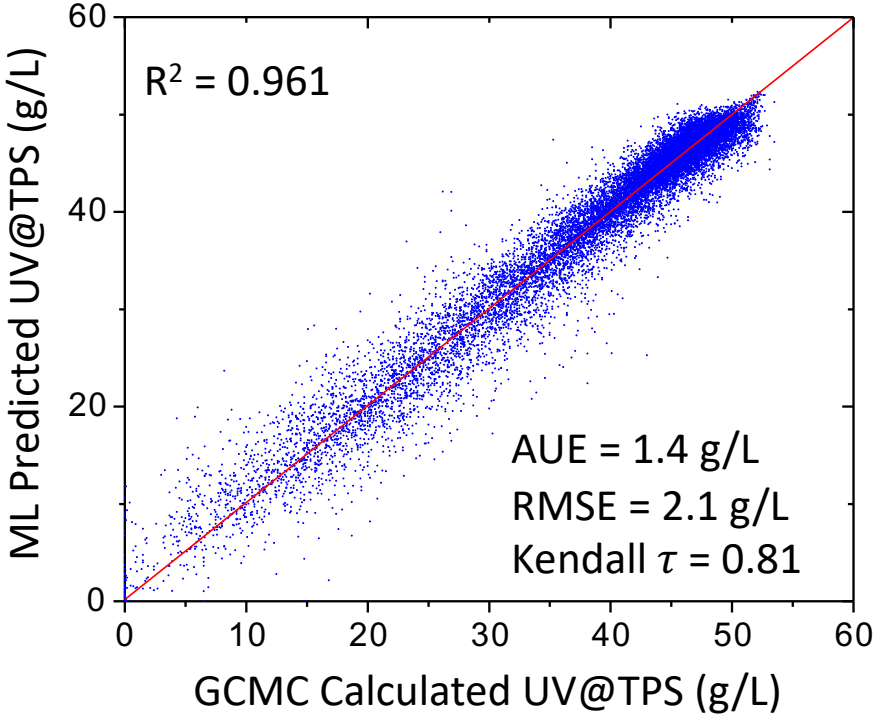
TPS: Benchmarking ML Predictions (1)

The Extremely Randomized Trees method is the best performing ML algorithm

Gravimetric Capacity



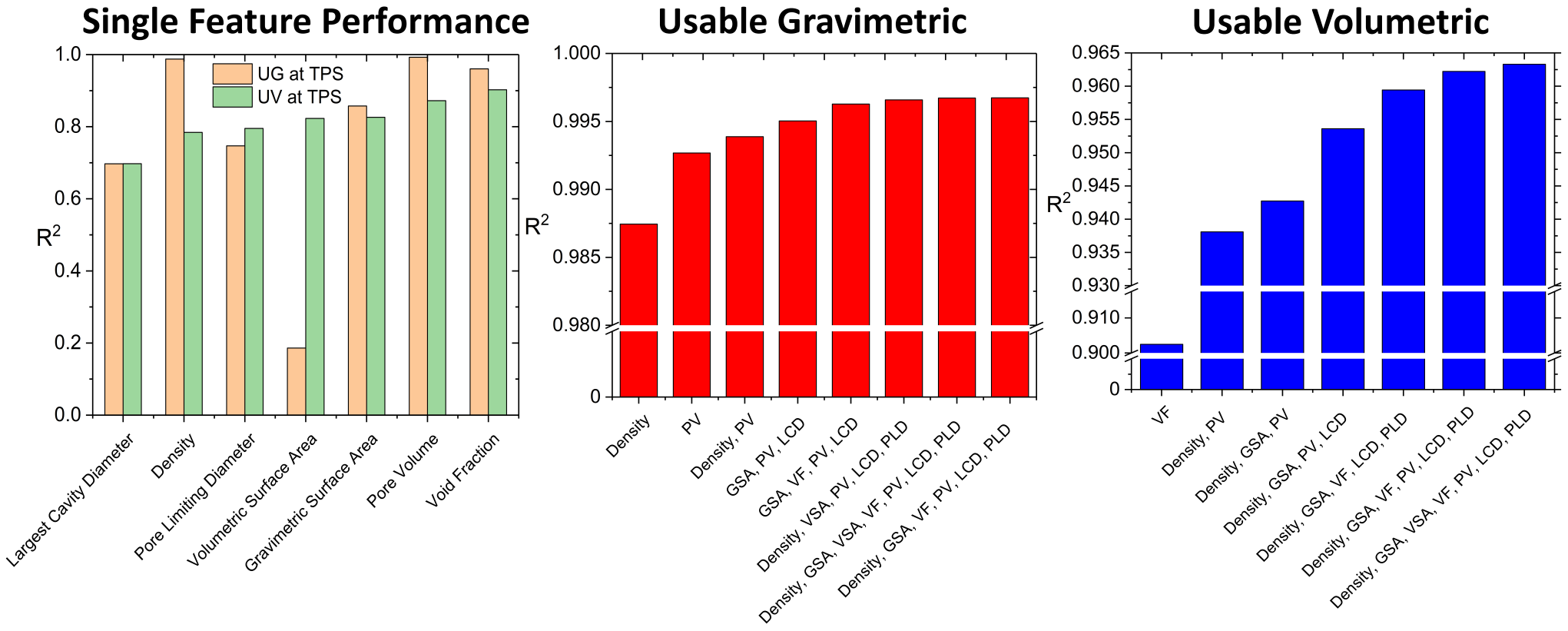
Volumetric Capacity



Correlations between ML predicted and GCMC calculated usable capacities of 24,741 MOFs (not used in training ML models) at 77 K for the temperature+pressure swing between 100bar/77K and 5bar/160K. ML models were developed by training Extremely Randomized Trees algorithm on a dataset of 74, 221MOFs.

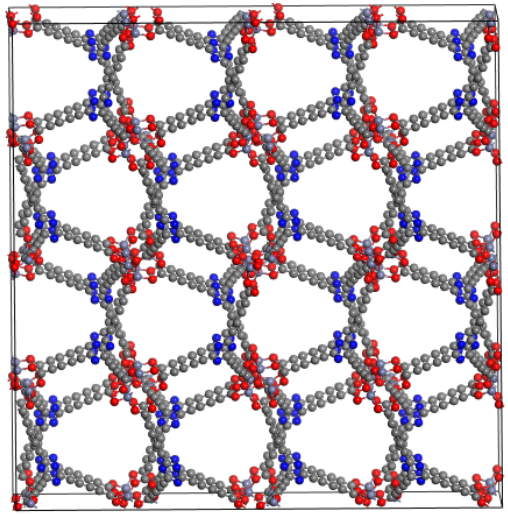
TPS: Benchmarking ML Predictions (2)

- Pore volume is the dominant feature that controls the usable volumetric capacity.
- Density and pore volume could be used to predict UV with ~ 94% accuracy.

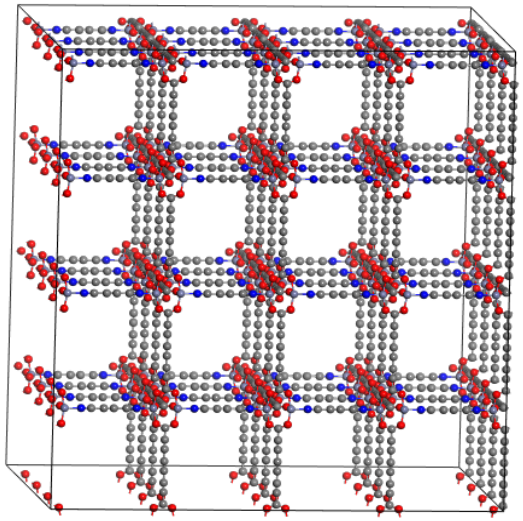


ML models were developed for all possible (2^7-1 = 127) combinations of features to identify the optimal feature set. Each histogram represents the highest R^2 value among all possible combinations of a given number of features.

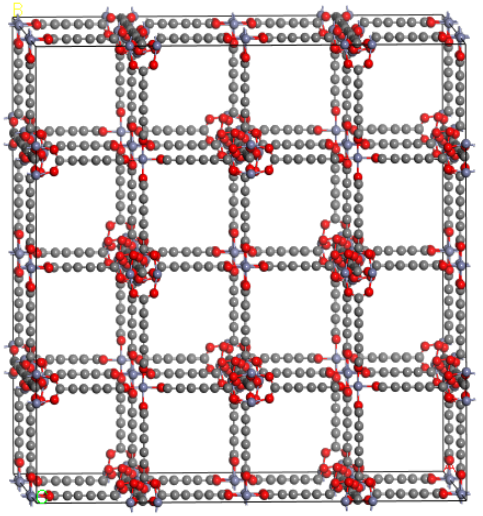
PS: Top 3 MOFs Visualized



Density: 0.30 g/cm^3
Gravimetric Surface Area: 5561 m^2/g
Volumetric Surface Area: 1695 m^2/cm^3
Pore Volume: 2.93 cm^3/g
Pore Diameter: 12.8 Å
Void Fraction: 0.89



Density: 0.31 g/cm^3
Pore Volume: 2.87 cm^3/g
Gravimetric Surface Area: 5926 m^2/g
Volumetric Surface Area: 1820 m^2/cm^3
Pore Diameter: 16 Å
Void Fraction: 0.88



Density: 0.31 g/cm^3
Pore Volume: 2.88 cm^3/g
Gravimetric Surface Area: 5073 m^2/g
Volumetric Surface Area: 1583 m^2/cm^3
Pore Diameter: 17.7 Å
Void Fraction: 0.90

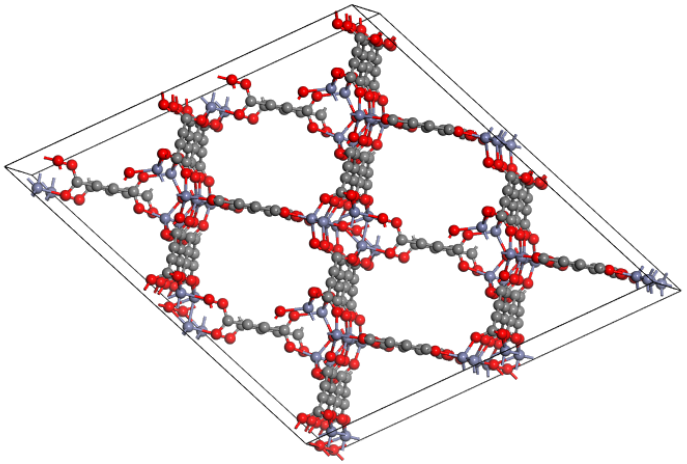
TPS High-Capacity MOFs synthesis (1)

Top performing MOFs according to our ML+GCMC calculations at temperature+pressure swing (TPS) conditions (between 100bar/77K and 5bar/160K).

Name	Source	Density (g/cm³)	Grav. Surf. Area (m²/g)	Vol. Surf. Area (m²/cm³)	Void Fraction	Pore Volume (cm³/g)	Largest Cav. Diam. (Å)	Pore Lim. Diam. (Å)	Usab. Grav. Cap. (wt.%)	Usab. Vol. Cap. (g/L)
MOF-5_cooh_2_16_4_basic_opt	Mail-order	0.70	3072	2154	0.68	0.68	7.8	12.2	8.0	61.1
MOF-5_cooh_2_2738_1_basic_opt	Mail-order	0.47	4548	2149	0.78	1.34	7.8	15.8	10.8	57.7
BOQQAB (MOF-650)	CSD refcode	0.49	3908	1919	0.85	1.73	18.3	9.9	10.2	56.5
MOF-5_cooh_2_972_1_basic_opt	Mail-order	0.67	3038	2037	0.74	0.95	6.7	11.9	7.5	54.9
hypotheticalMOF_5056615_i_1_j_29_k_2_m_2_cat_1	Northwesteren	0.56	4388	2474	0.79	1.41	7.9	9.6	8.6	53.8
ODIXEG (PCN-516)	CSD refcode	0.55	4090	2259	0.84	1.42	10.4	7.5	8.8	53.7
hypotheticalMOF_5057692_i_1_j_29_k_19_m_2	Northwesteren	0.55	4546	2489	0.80	1.47	7.2	9.4	8.8	53.6
ENITAX	CSD refcode	0.57	4021	2304	0.83	1.36	10.1	7.2	8.5	53.5
FINJAO	CSD refcode	0.47	6977	3258	0.80	1.70	7.4	6.4	10.2	53.5
TEQPEM	CSD refcode	0.57	3456	1980	0.86	1.45	17.2	9.2	8.5	53.5
MOF-5									**7.8**	**51.9**

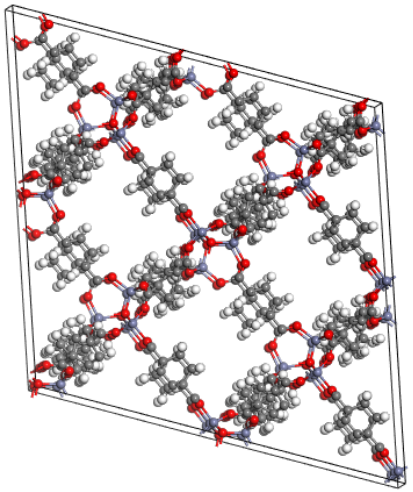
From these MOFs, red highlighted MOFs were chosen for attempted synthesize and activation

TPS High-Capacity MOFs synthesis (2)



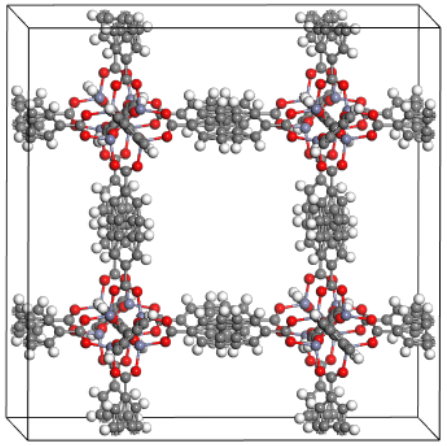
MOF-5_cooh_2_16_4_basic_opt
(mail-order)

Density: 0.70 g/cm^3
Pore Volume: 0.68 cm^3/g
Gravimetric Surface Area: 3072 m^2/g
Volumetric Surface Area: 2154 m^2/cm^3
Pore Diameter: 7.8 Å
Void Fraction: 0.68



MOF-5_cooh_2_972_1_basic_opt
(mail-order, TMOF-2)

Density: 0.67 g/cm^3
Pore Volume: 0.95 cm^3/g
Gravimetric Surface Area: 3038 m^2/g
Volumetric Surface Area: 2037 m^2/cm^3
Pore Diameter: 6.7 Å
Void Fraction: 0.74



BOQQAB/MOF-650

Density: 0.49 g/cm^3
Pore Volume: 1.73 cm^3/g
Gravimetric Surface Area: 3908 m^2/g
Volumetric Surface Area: 1919 m^2/cm^3
Pore Diameter: 18.3 Å
Void Fraction: 0.85

Red highlighted MOFs were chosen for attempted synthesize and activation

Reviewer-Only Slides

Data Management Plan

This data management plan (DMP) explains how data generated during work performed for EERE will be shared and preserved in accordance with DOE data management regulations. Based on the performance record from past DOE awards, the PIs have followed DOE's policy of prompt publication of sponsored research results and have communicated the results at conferences, meetings, and on our web sites.

1. Data Types and Sources

As described in the main body of the proposal, this project involves simulations and characterization of hydrogen storage materials. The data to be produced will include information on models, input parameters, model predictions, and data collected from experiments. The project will not involve animal or human subjects' data.

2. Data Content and Format

The raw data from the and experimental measurements are typically a series of arrays of two or three dimensions. These data will be visualized, which will yield images (in a common format, such as tiff and jpg), and will be analyzed to extract features such as performance trends. These data will be stored with associated metadata containing the model information, input parameters, and date/time of the experiment. Essential metadata will be stored in an electronic form as well as dated and indexed laboratory notebooks. Modeling and data analysis codes are typically written in FORTRAN, C++, MATLAB, and IDL.

3. Sharing and Preservation

The analyses of the modeling and experimental data (graphs, tables, equations) are to be published in peer-reviewed journal articles, conference proceedings, and book chapters. The data used to generate these analyses will be contributed as Electronic Supporting Information to each journal publication. Additional unpublished data deemed useful to the community will be made publically available. Some data will be made available online, after formal publication, via the University of Michigan's Deep Blue Data service (https://deepblue.lib.umich.edu/data) or the PIs University web page. Data will also be made available to interested parties upon request, consistent with scientific journal access privileges. If application for intellectual property rights is expected based on parts of the data developed in this project, access to the data will be granted upon request once provisional patent filings are made, but no longer than one year after filing an invention disclosure.

Data will be stored for at least three years after completion of the project, per NSF rules. Published data will also be available in print or electronically from publishers for much longer, subject to their access charges. Theses are retained in databases and hard copies in the PI's office indefinitely. All electronic data will also be stored in laboratory computers and a backup hard drive located in our laboratories. Original laboratory notebooks will be kept by the PI in his office or laboratory. Upon changes in project personnel, such as student graduation, the PI will take over responsibility for the data.

4. Protection

We do not anticipate privacy or security issues with the data generated in this project. More specifically, the generated data will not contain any personally identifiable information, nor contain any business confidential information. Efforts will be undertaken to ensure that the data has no negative impact on innovation, or on US competitiveness.

5. Rationale

The data generated in this project will be of value to scientists conducting research in metal-organic frameworks (MOFs), and to engineers exploring the use of MOFs in gas capture and storage applications. This project may suggest new MOF designs with improved performance, which could positively impact the efficiency of vehicle transportation.