Optimizing MOF properties for seasonal heat storage: a machine learning approach

Original
Optimizing MOF properties for seasonal heat storage: a machine learning approach / Trezza, G., Bergamasco, L., Fasano, M., Chiavazzo, E.. - In: JOURNAL OF PHYSICS. CONFERENCE SERIES. - ISSN 1742-6588. - 2766:(2024). (Eurotherm 2024 Bled, Slovenia) [10.1088/1742-6596/2766/1/012219].

Publisher:
IOP Publishing Ltd

Published
DOI:10.1088/1742-6596/2766/1/012219

Journal of Physics: Conference Series

Optimizing MOF properties for seasonal heat storage: a machine learning approach

To cite this article: G Trezza *et al* 2024 *J. Phys.: Conf. Ser.* **2766** 012219

9th European Thermal Sciences Conference (Eurotherm 2024) IOP Publishing
Journal of Physics: Conference Series **2766** (2024) 012219 doi:10.1088/1742-6596/2766/1/012219

Optimizing MOF properties for seasonal heat storage: a machine learning approach

G Trezza, L Bergamasco, M Fasano and E Chiavazzo
Department of Energy, Politecnico di Torino, C.so Duca degli Abruzzi 24, Torino 10129, Italy
E-mail: `eliodoro.chiavazzo@polito.it`

Abstract. In the quest to enhance thermochemical energy storage using promising sorbents, this work presents a study on the optimization of Metal Organic Frameworks (MOFs) properties for gas sorption, with a focus on CO_2 and H_2O adsorption. Through the analysis of crystallographic descriptors, the study aims to streamline the selection of MOFs that could potentially exceed the performance of existing water sorbent pairs. A comprehensive comparison of sequential learning (SL) algorithms reveals a method for identifying the minimal set of descriptors that influence adsorption properties of MOFs. The protocol involves constructing and training machine learning (ML) models to determine the number of influential descriptors and utilizing SHAP analysis to evaluate their importance. Findings suggest that including only these critical descriptors in the exploration space reduces computational load. Notably, the COMBO and the FUELS algorithms consistently outshine random guessing, validating their efficacy in materials optimization. The challenge of accessing full adsorption properties across the entire coverage range is addressed by a computational screening procedure requiring minimal input data. This method suggests that some vanadium based MOFs, originally designed for different purposes, could surpass the current leading compounds for thermal energy storage, primarily due to their optimal Henry coefficient values for water adsorption.

1. Introduction

Metal-organic frameworks (MOFs) are crystalline structures comprising metal ions and organic linkers, characterized by tunable porosity and remarkably large surface area [1]. Such unique properties make MOFs suitable porous adsorbents for capturing CO_2, potentially contributing to the mitigation of greenhouse gas emissions [2]. However, in the realm of engineering applications, optimization challenges arise from the presence of diverse inlet gas streams, different operating conditions, and customized target properties for each specific scenario. This variability makes complex to pinpoint an ideal MOF crystal that suits all applications [3], leading to a case-by-case optimization. In this regard, modern sequential learning (SL) algorithms are becoming notably effective for navigating into the high-dimensional (crystallographic) feature space of materials. Specifically, when assessing a complex black-box function by means of resource-intensive physical or numerical experiments, SL tools offer a systematic approach to smartly explore the high-dimensional parameter (feature) space [4, 5, 6].

The primary goal of this study [7] is to identify the essential set of features (or descriptors [8]) governing critical MOFs adsorption properties in the low-coverage regime. Specifically, this pertains to the Henry solubility coefficients for both CO_2-MOFs and H_2O-MOFs working pairs. This minimal set encompasses crucial crystallographic features that underlie a specific adsorption property of interest. In this context, each minimal set of descriptors can be considered the *genetic code* for a particular property, and the identification process is outlined below.

First, we take advantage of a database made of 8206 hypothetical MOFs [9], featurizing the corresponding Crystallographic Information Files (CIFs) with 1557 Classical Force fields Inspired Descriptors (CFID) [10]. Subsequently, we train and validate regression models for predicting key properties in heat storage applications [11, 12]. We thus rank and select the minimal set of such descriptors by assessing the importance of each feature over the corresponding model outputs via the Tree SHAP interpretation algorithm [13], which is indeed widely used in Materials Science [14, 15, 16]. Moreover, we investigate the role of those descriptors in SL optimization over several strategies from three SL different methodologies: (a) Random Forests with Uncertainty Estimates for Learning Sequentially (FUELS) [17]; (b) kriging algorithm [18]; (c) COMmon Bayesian Optimization Library (COMBO) [19]. Also, we evaluate the performance of SL optimization using both the minimal subset of features (derived from the pipeline and SHAP analysis) and a more extensive set of variables. This comparison aims to reveal the impact of descriptor identification on the minimum number of experiments required to identify a MOF with the highest value of the desired property. Finally, engineering figures of merit optimization requires getting access to the material characterization in a wide coverage range. However, often only the low coverage regime (i.e., Henry coefficient) is known. Therefore, we propose a fast procedure assessing a crucial performance metric in closed water-sorption seasonal thermal energy storage applications, i.e., the specific stored energy.

2. Descriptors of sorption properties in MOFs and their use in SL algorithms

In the database by Boyd *et al.* [9], simulated sorption properties are available for over 8000 hypothetical MOFs, together with the corresponding CIFs. We thus generated four datasets, each encompassing the same (both chemical and structural) 1557 Matminer-based [20] features, and a target property among Henry coefficient for CO_2 (8194 entries), working capacity for CO_2 (8202 entries), Henry coefficient for H_2O (8202 entries), and surface area (5028 entries). The variation in the number of entries is attributed to missing values for certain properties in ref. [9]. Based on such data and using a Random Forest-based pipeline with hyperparameter tuning in a five-fold cross-validation, we trained four distinct ML models to predict the four aforementioned properties.

Via the TreeSHAP routine [21, 13], we identified the important features as those contributing to the 75% of the cumulative curves over the importance coefficients, assessing the performance of SL algorithms for the targeted sorption properties. This evaluation involved comparing the use of two sets of descriptors: (a) the reduced set of important descriptors and (b) a larger set comprising 100 descriptors, which includes the aforementioned important ones and some additional (non-meaningful) ones. Surprisingly, optimization with SL in the space of relevant descriptors does not consistently guarantee a quicker convergence of the procedure toward the optimum property value. Moreover, among the three examined methodologies, both COMBO and FUELS methods consistently demonstrated faster convergence to the optimum value compared to the random choice strategy.

3. Optimization under incomplete access to the isosteric field of candidate MOFs-water working pairs

Our objective is to assess the performance of hypothetical MOFs for a significant energy engineering application, i.e., seasonal thermal energy storage through water sorption. To this end, we propose a methodology to obtain the complete isosteric field for each hypothetical MOF-water pair, whereas only the Henry low-coverage regime is available; by taking advantage of such procedure, we rank the hypothetical MOFs by Boyd *et al.* [9] in terms of the specific stored energy over a seasonal cycle.

Specifically, having in mind space heating applications in temperate climates, we define a thermodynamic ideal cycle bounded by four operating temperatures: $T_\mathrm{A} = 308\,\mathrm{K}$ (the minimum temperature on the user side), $T_\mathrm{C} = 353\,\mathrm{K}$ (the maximum temperature on the source side), $T_\mathrm{E} = 278\,\mathrm{K}$ (the average winter temperature), $T_\mathrm{F} = 303\,\mathrm{K}$ (the average summer temperature). The equilibrium water vapor pressures $p_\mathrm{E} = 866.2\,\mathrm{Pa}$ and $p_\mathrm{F} = 4231.6\,\mathrm{Pa}$ at evaporator and condenser respectively are also defined considering the Antoine equation at the average winter and summer temperatures.

In particular, the adsorption/desorption of an infinitesimal amount of adsorbate moles $\mathrm{d}n$ is possible upon exchanging an infinitesimal heat $\mathrm{d}Q = q_\mathrm{st}\mathrm{d}n$, where q_st denotes the isosteric heat. It can be demonstrated that the specific stored energy over the cycle is approximately $n_\mathrm{TOT} q_\mathrm{st} \Delta\theta$, with n_TOT denoting the available adsorption sites per unit of dry adsorbent mass, and $\Delta\theta$ the coverage span over the cycle. We report here below the primary simplifying assumptions of our approach:

- The quantity n_TOT is closely related to the water uptake, which is not available from the published data in ref. [9]. However, by means of data published by Chaemchuen *et al.* [22], we found a good correlation between the water uptake and the internal surface area, yielding the relation water uptake $= \eta \times$ surface area, with $\eta = 3.875 \times 10^{-4}\,\mathrm{g_{H_2O}\,m^{-2}}$.
- To draw the adsorption isotherms over the entire coverage regime we adopted the Frumkin–Fawler–Guggenheim (FFG) model, which is suitable for the *S-shaped* isotherms typically observed in MOFs/water systems. Such model relies upon only the Henry coefficient $H(T)$ in units of Pa^{-1} at a certain temperature T. However, Boyd *et al.* [9] made available such Henry coefficients $\widetilde{H}(T_0)$ in units of $\mathrm{mol_{H_2O}\,kg^{-1}_{MOF}\,bar^{-1}}$ at the reference temperature $T_0 = 298\,\mathrm{K}$. We demonstrate [7] that a proper conversion formula yields $H(T_0) = \widetilde{H}(T_0)\mathcal{M}_\mathrm{H_2O}/(\eta S) \times 10^{-8}$, with S being the internal surface area, $\mathcal{M}_\mathrm{H_2O}$ the molecular weight of water and where the factor 10^{-8} appears because $[\mathcal{M}_\mathrm{H_2O}] = \mathrm{g_{H_2O}\,mol^{-1}_{H2O}}$, $[\eta] = \mathrm{g_{H_2O}\,m^{-2}}$, $[S] = \mathrm{m^2\,g^{-1}_{MOF}}$, and so $[\widetilde{H}(T_0)\mathcal{M}_\mathrm{H_2O}/(\eta S)] = \mathrm{g_{MOF}\,kg^{-1}_{MOF}\,bar^{-1}}$. Also, in order to reconstruct adsorption isotherms at a generic temperature T, we adopted the Polanyi potential theory. Indeed, the Polanyi potential is defined as $\mathcal{A} = -RT\ln(p_s(T)/p)$, where $R = 8.314\,\mathrm{Jmol^{-1}K^{-1}}$, $p_s(T)$ denotes the saturation pressure of water at temperature T, while p is the pressure of the vapor phase on the adsorbent surface [23]. From the practical standpoint, since $\mathcal{A}$ is a constant of the sorption pair for a specific pressure p, we calculated $\mathcal{A}$ at T_0; subsequently, we adjusted the abscissa p of the isotherm obtained at temperature T_0 by scaling it according to $p = p_s(T)\exp(\mathcal{A}/(RT))$.
- We computed the isosteric heat by means of the Clausius-Clapeyron equation, yielding:

$$q_\mathrm{st} = \frac{R}{3}\frac{T_\mathrm{C}T_\mathrm{A}}{T_\mathrm{C} - T_\mathrm{A}}\sum_{i=1}^{3}\ln\frac{p_2(\theta_i)}{p_1(\theta_i)}, \tag{1}$$

  where points 1 and 2 denote the intersections of an isosteric transformation with the respective isotherms at temperatures T_A and T_C; we reiterated the methodology for three coverage values ($\theta_1 = 0.4$, $\theta_2 = 0.5$, $\theta_3 = 0.6$), getting the average.
- Ultimately, after determining the two isotherms at T_A and T_C, we derived the coverage span $\Delta\theta$ based on the thermodynamic ideal cycle within the $\theta - p$ chart.

The whole procedure above is summarized in Fig. 1. We thus calculated the objective function $Sq_\mathrm{st}\Delta\theta$, representing the specific stored energy up to a constant, for the complete set of 5028 hypothetical MOFs characterized by a positive surface area in ref. [9]. Notably, the

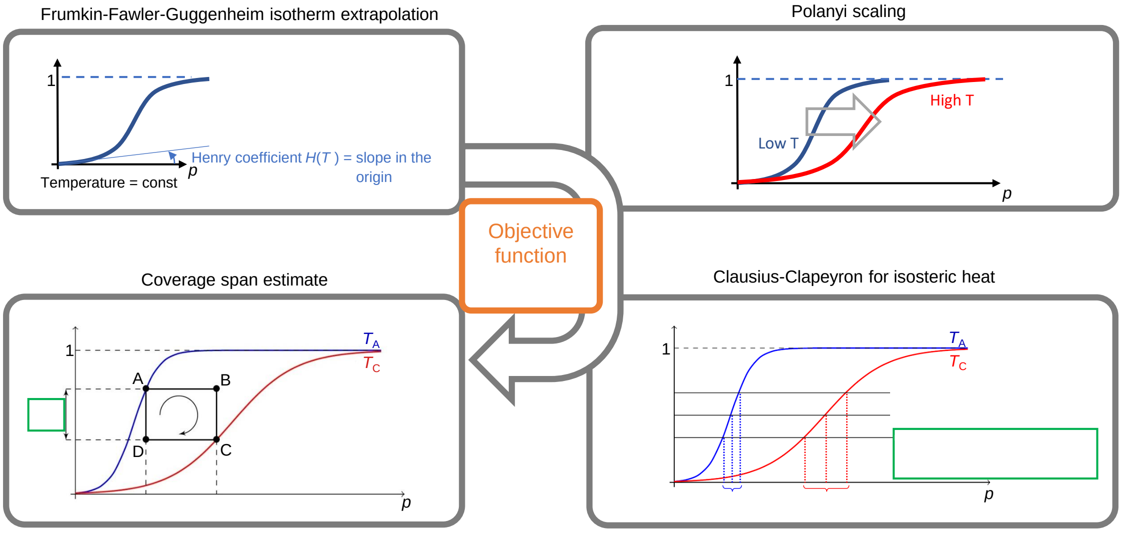

Figure 1. Proposed methodology for determining the specific energy of MOF-water systems when only partial information about the isosteric field is available. By utilizing the Henry coefficient for H_2O obtained from literature at a specific temperature, an isotherm is generated using the Frumkin–Fawler–Guggenheim model. The Polanyi potential is employed for temperature scaling. Once two relevant isotherms are identified and upon definition of the necessary environmental conditions, the specific stored energy can be computed. This involves calculating the isosteric heat q_{st} and determining the coverage span $\Delta\theta$. In this study, $\beta = 3.4$ always (for details, see Supplementary Note 6 of ref. [7]).

top-performing potential MOFs, when ranked in terms of the specific energy, are all Vanadium-based. This can be attributed to the Henry coefficient values for H_2O falling within the optimal range, resulting in a favorable coverage span $\Delta\theta$ over the thermodynamic cycle. Among those, the four top-performing are predicted to exhibit (material-based) specific energy values ranked among the highest reported in the literature for sorption-based thermal energy storage under similar operating conditions, as showcased in Fig. 2.

4. Conclusion

In this study we demonstrated that a general procedure for identifying the essential set of governing descriptors for a specific adsorption property involves two key steps: i) developing and training a ML model to detect the relevant descriptors; ii) assessing the relative importance of each explanatory variable on the selected output by means of the SHAP analysis. We obtained that incorporating the set of essential descriptors (pertaining to a specific property of interest) within the feature space does not impact the SL convergence performance, even though its computational load is influenced by the dimension of the parameter space under investigation. Focusing solely on the most significant features may indeed be advantageous in managing computational burden. Moreover, we noticed that the COMBO and FUELS algorithms consistently perform better than random guessing.

Importantly, we proposed a fast procedure for assessing the performances of water-MOF pairs according to an engineering figure of merit for thermal seasonal application (i.e., the specific stored energy), without having explicitly access to the entire coverage regime. Such methodology suggests that, in this regard, top performing hypothetical MOFs in the database published by Boyd *et al.* [9] are all Vanadium based, and are worth of further investigation.

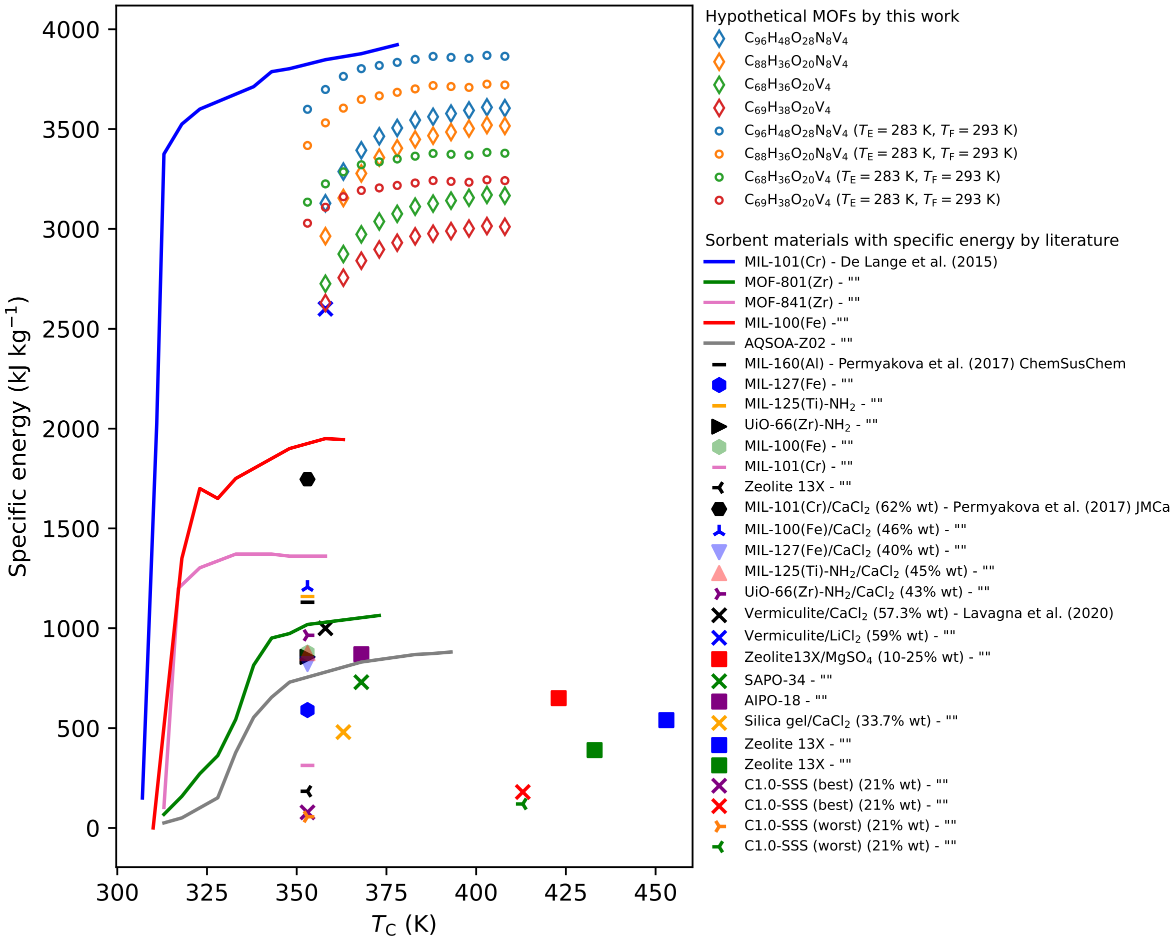

Figure 2. Comparison of the expected specific energy for various desorption temperatures, denoted as T_C, among the optimal MOFs identified in this study, with various water adsorbent materials found in the literature [24, 25, 26, 27]. We consider either standard environmental conditions, i.e. evaporation temperature $T_E = 278\,K$ and condensation temperature $T_F = 303\,K$, or conditions of $T_E = 283\,K$ and $T_F = 293\,K$, adsorption temperature $T_A = 308\,K$ always. This figure is taken from [7].

5. Acknowledgments

E.C. acknowledges financial support of the Italian National Project PRIN Heat transfer and Thermal Energy Storage Enhancement by Foams and Nanoparticles (2017F7KZWS) and of the research contract PTR 2019/21 ENEA (Sviluppo di modelli per la caratterizzazione delle proprietà di scambio termico di PCM in presenza di additivi per il miglioramento dello scambio termico) funded by the Italian Ministry of Economic Development (MiSE).

References

[1] Kitagawa S *et al.* 2014 *Chem. Soc. Rev.* **43** 5415–5418
[2] Chen S, Lucier B E, Boyle P D and Huang Y 2016 *Chem. Mater.* **28** 5829–5846
[3] Ejeian M and Wang R 2021 *Joule* **5** 1678–1703
[4] Liang Q, Gongora A E, Ren Z, Tiihonen A, Liu Z, Sun S, Deneault J R, Bash D, Mekki-Berrada F, Khan S A *et al.* 2021 *Npj Comput. Mater.* **7** 1–10
[5] Kim Y, Kim E, Antono E, Meredig B and Ling J 2020 *Npj Comput. Mater.* **6** 1–9

[6] Bonke S, Trezza G, Bergamasco L, Song H, Rodríguez-Jiménez S, Hammarström L, Chiavazzo E and Reisner E 2023
[7] Trezza G, Bergamasco L, Fasano M and Chiavazzo E 2022 *npj Computational Materials* **8** 123
[8] Gomes S I, Amorim M J, Pokhrel S, Mädler L, Fasano M, Chiavazzo E, Asinari P, Jänes J, Tämm K, Burk J *et al.* 2021 *Nanoscale* **13** 14666–14678
[9] Boyd P G, Chidambaram A, García-Díez E, Ireland C P, Daff T D, Bounds R, Gładysiak A, Schouwink P, Moosavi S M, Maroto-Valer M M *et al.* 2019 *Nature* **576** 253–256
[10] Choudhary K, DeCost B and Tavazza F 2018 *Phys. Rev. Mater.* **2** 083801
[11] Fasano M, Borri D, Chiavazzo E and Asinari P 2016 *Appl. Therm. Eng.* **101** 762–769
[12] Fasano M, Bergamasco L, Lombardo A, Zanini M, Chiavazzo E and Asinari P 2019 *Front. Energy Res.* **7** 148
[13] Lundberg S M, Erion G, Chen H, DeGrave A, Prutkin J M, Nair B, Katz R, Himmelfarb J, Bansal N and Lee S I 2020 *Nat. Mach. Intell.* **2** 2522–5839
[14] Tang B, Lu Y, Zhou J, Chouhan T, Wang H, Golani P, Xu M, Xu Q, Guan C and Liu Z 2020 *Materials Today* **41** 72–80
[15] Zhong X, Gallagher B, Liu S, Kailkhura B, Hiszpanski A and Han T Y J 2022 *npj Computational Materials* **8** 204
[16] Trezza G and Chiavazzo E 2023 *Materials Today Communications* **36** 106579
[17] Ling J, Hutchinson M, Antono E, Paradiso S and Meredig B 2017 *Integr. Mater. Manuf. Innov.* **6** 207–217
[18] Lophaven S N, Nielsen H B, Sondergaard J and Dace A 2002 *Technical University of Denmark, Kongens Lyngby, Technical Report No. IMMTR-2002* **12**
[19] Ueno T, Rhone T D, Hou Z, Mizoguchi T and Tsuda K 2016 *Mater. Discov.* **4** 18–21
[20] Ward L, Dunn A, Faghaninia A, Zimmermann N E, Bajaj S, Wang Q, Montoya J, Chen J, Bystrom K, Dylla M *et al.* 2018 *Comput. Mater. Sci.* **152** 60–69
[21] Lundberg S M and Lee S I 2017 *Adv. Neural Inf. Process. Syst.* **30**
[22] Chaemchuen S, Xiao X, Klomkliang N, Yusubov M S and Verpoort F 2018 *Nanomaterials* **8** 661
[23] Butt H J, Graf K and Kappl M 2013 *Physics and chemistry of interfaces* (John Wiley & Sons)
[24] de Lange M F, Verouden K J, Vlugt T J, Gascon J and Kapteijn F 2015 *Chem. Rev.* **115** 12205–12250
[25] Permyakova A, Skrylnyk O, Courbon E, Affram M, Wang S, Lee U H, Valekar A H, Nouar F, Mouchaham G, Devic T *et al.* 2017 *ChemSusChem* **10** 1419–1426
[26] Permyakova A, Wang S, Courbon E, Nouar F, Heymans N, d'Ans P, Barrier N, Billemont P, De Weireld G, Steunou N *et al.* 2017 *J. Mater. Chem. A* **5** 12889–12898
[27] Lavagna L, Burlon D, Nisticò R, Brancato V, Frazzica A, Pavese M and Chiavazzo E 2020 *Sci. Rep.* **10** 1–13