Journal of Materials Chemistry A

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d5ta10538b

Application of multi-task learning in analysing the methane working capacity of metal–organic frameworks

Junhui Kou, [abc] Tianle Liu, [*ab] Guosheng Jiang,[ab] Guokun Yang,[ab] Zerang Li,[ab] Xiaoyang Ni[b] and Hao Liu[b]

The screening and rational design of metal–organic frameworks (MOFs) with optimal methane capture performance remain a critical challenge for environmental and energy applications. Existing research often emphasizes single-point adsorption capacity, overlooking the dynamic mechanisms relevant to pressure swing adsorption (PSA) required for practical use. To address these challenges, this study developed a multi-task learning (MTL) framework for high-throughput screening of methane adsorption in MOFs. The proposed model integrates attention and screening mechanisms to predict the methane working capacity using a dataset of 252 352 MOFs characterized by geometric and chemical descriptors. Shared parameters were leveraged across adsorption prediction tasks under six pressure conditions, achieving high predictive accuracy, with R^2 values of 0.992 and 0.962 for gravimetric and volumetric working capacities, respectively. SHAP analysis within the MTL context identified shared underlying mechanisms governing adsorption at both low and high pressures. Key descriptors influencing adsorption capacity included accessible pore volume and specific surface area, alongside chemically relevant atomic identity, covalent radius, and nuclear charge. Variations in feature importance from low to high pressures reflected the shifts in methane adsorption mechanisms. This MTL model provides a novel approach to accelerate the discovery of MOFs with high working capacities, offering new insights into methane adsorption mechanisms in MOFs under varied pressure regimes.

Received 27th December 2025
Accepted 27th February 2026

DOI: 10.1039/d5ta10538b

rsc.li/materials-a

1. Introduction

Methane is a potent greenhouse gas.[1] Anthropogenic emissions have risen rapidly over recent decades, posing a significant threat to global climate and causing widespread environmental issues.[2,3] Efficient capture, purification, separation, and storage of methane are essential for mitigating its environmental impact.[4] Metal–organic frameworks (MOFs), characterized by high surface areas and large pore volumes,[5] are among the most promising materials for methane capture.[6] MOFs possess low density, high porosity, abundant adsorption sites,[7] as well as diverse and highly ordered porous structures[8,9] combined with tunable chemical functionalities.[10] These properties make MOFs suitable for methane capture, storage, and purification applications.[11] However, the vast structural diversity and complexity of MOFs pose significant challenges to the rational design and synthesis of materials with optimized adsorption performance.[12,13] The methane adsorption process is a critical component of methane capture, storage, and desorption, and understanding this process is fundamental for designing superior adsorbent materials.[14] Studies have shown that the pore structure, surface characteristics, properties of adsorption sites, and interaction patterns between MOFs and methane significantly influence the adsorption behavior.[15–17] Therefore, investigating the methane adsorption process is essential for designing effective methane adsorbents and enhancing the practical performance of MOFs.

Enhancing the structural design, functional groups, and adsorption sites of MOFs to develop materials with increased methane uptake has become a central focus of current research.[18,19] Numerous studies[20,21] have aimed to push the limits of methane adsorption in MOFs. Gándara *et al.* employed a strategy of dangling BTB ligands to tailor the pore structure of MOFs, reducing the pore size to 7.6 Å, and synthesized MOF-519, which exhibited a record methane uptake of 279 cm^3 cm^{-3} at 80 bar.[22] This work demonstrated the critical role of optimizing pore structure for enhanced methane adsorption. These successful studies highlight that understanding the key material properties and their synergistic effects during adsorption is essential to advancing MOFs' methane uptake performance.[23] However, experimentally investigating the individual and combined effects of various factors on methane

[a]Unconventional Cementing and Special Reinforcement Laboratory, China University of Geosciences, Wuhan 430074, China. E-mail: liutianle@cug.edu.cn
[b]Faculty of Engineering, China University of Geosciences, Wuhan, 430074, China
[c]Department of Applied Science and Technology, Politecnico di Torino, Duca degli Abruzzi 24, Torino 10129, Italy

adsorption remains a complex process that requires substantial time and resource.[24] Compared to traditional experimental methods based on experience and trial-and-error, machine learning (ML) provides a more efficient way to accelerate research progress.[19,25,26] Fernandez *et al.* applied classical ML models to identify high-performance MOFs for carbon capture and separation.[27,28] The application of ML relies on material descriptors converted from chemical or physical information as input features, which are the key to interpret the resulting predictions,[29] to train predictive models. As research progressed, the variety of MOF descriptors has expanded to hundreds or even thousands, and this large number of feature types may complicate model training and optimization.[30] Selecting appropriate descriptors is key to reducing bias, ensuring generalization, and improving research efficiency in ML models.[31] The development of interpretable ML techniques holds promise to address this challenge.[32] Numerous studies have employed Shapley additive explanations (SHAP)[33] to conduct interpretable analyses of MOFs' structure–property relationships.[30,34,35] Wei *et al.* utilized permutation feature importance and transfer learning to investigate methane adsorption in MOFs, highlighting the significance of methane-related descriptors under high-pressure conditions.[36] Interpretable ML bridges hypothetical MOF data with experimental, reducing the experimental time and costs.[25,37] In summary, advances in ML currently provide efficient assistance for MOF design and performance prediction, but challenges remain, including training data quality and model optimization, which require further research.

Although extensive research has focused on maximizing the methane adsorption capacity, pressure swing adsorption (PSA) represents the key technique used in industrial applications.[38–40] Materials with effective working capacity are characterized by a larger difference in methane loading between high-pressure adsorption and low-pressure desorption.[41] However, current high-throughput screening and predictive models predominantly focus on single-point adsorption capacities,[42,43] with few studies addressing the dynamic characteristic of the PSA process. Since PSA involves adsorption at high pressure and desorption at low pressure, understanding the mechanisms and characteristics of MOFs under these conditions is essential for designing materials with high working capacities. In fact, conventional molecular simulations and ML or neural network models cannot effectively predict this dynamic process. This limitation impedes the identification of MOFs suitable for practical methane storage applications. To address this issue, several researchers have attempted to predict methane adsorption isotherms of MOFs. For example, existing studies used multilayer perceptron models and deep feed-forward neural networks to train methane uptake data at six different pressures and thereby fit isotherms.[44,45] However, these approaches treat adsorption predictions at each pressure as separate tasks, overlooking the inherent correlations among adsorption behaviors across different pressures. This task separation not only results in inefficient use of computational resources but also restricts the model's ability to capture the underlying shared mechanisms between varying pressure adsorption, thereby diminishing the prediction accuracy and generalization performance.

In this study, the multi-task learning (MTL) approach is intended to address the challenges described above. As an effective deep learning paradigm, MTL enables simultaneous learning of multiple related tasks within a single model, facilitating information sharing and collaborative optimization.[46] For example, some researchers developed an MTL model to selectively adsorb oils and phenolic pollutants from coal chemical wastewater, simultaneously predicting adsorption rates and capacities, and revealing co-adsorption mechanisms in complex systems.[47] Other researchers designed an MTL model based on soft parameter sharing to address the dual challenge of cadmium contamination and methane emissions in soil remediation using biochar, achieving superior prediction of their synergistic effects compared to traditional single-task models.[48] In the context of methane adsorption prediction, MTL treats adsorption capacities at different pressures as interconnected output tasks. By leveraging intrinsic relationships among these tasks, it captures both shared features and distinct mechanisms between high- and low-pressure adsorption processes. This approach enables a more comprehensive understanding of MOFs' adsorption behavior across varying pressures and improves the prediction accuracy.

To address the challenges in designing MOFs and to explore efficient methods for screening materials with practical methane working capacity, this study introduces MTL into the development of adsorption prediction models for the first time. The approach integrates the attention mechanism and working capacity screening mechanism to establish a novel predictive framework. Adsorption and working capacity prediction models were successfully developed, demonstrating superior performance under multiple pressure conditions compared to existing studies. The intrinsic relationships learned between low- and high-pressure adsorption tasks by the MTL model were thoroughly investigated and interpreted through SHAP analysis. The interactive contributions of individual and combined descriptors to MOF adsorption performance were systematically examined. By exploiting inherent correlations among adsorption properties, a more comprehensive and practical evaluation of MOF performance was achieved using this approach. This study enhanced the accuracy of methane working capacity predictions and introduced a novel method to accelerate the discovery of MOFs with practical application potential.

2. Computational methods

2.1 MOF dataset and descriptors

Crystallographic data of MOFs, including both experimentally synthesized and hypothetical structures, were extracted from eight public databases.[49] The dataset encompasses a wide variety of metal nodes and organic ligands. Duplicate structures were removed to eliminate redundancy. Further filtering was performed following Chung *et al.*[50] to exclude conformational isomers and functional group variants differing only in substitution positions. Chemical validity of MOF structures was verified using the method of Burner *et al.*,[51] leading to the removal

of structures containing improper oxidation states and over-coordinated main group elements. A comprehensive validity check was conducted across all MOFs using the MOSAEC approach introduced by White *et al.*,[52] enabling the elimination of frameworks with chemically unreasonable metal oxidation states and yielding a final dataset of 252 352 structures.

To construct the input feature space, both geometric and chemical descriptors were incorporated into the predictive models. Fourteen geometry-related parameters were extracted for each MOF using the Zeo++ software package,[53] employing a methane-sized probe radius of 1.86 Å.[51] These descriptors, which have been extensively validated in prior studies,[42,54] encompass metrics describing the pore size, accessible volume, surface area, and framework density. A complete list and definitions of the geometric descriptors are provided in Table S1.

Chemical descriptors were characterized using revised autocorrelation function (RAC) descriptors. The RAC methodology was originally introduced by Janet *et al.*[55] and subsequently extended by Moosavi *et al.*[56] to enable its application to periodic MOF systems. By encoding correlations among heuristic atomic properties within graph-based representations of MOFs, RAC descriptors provide an effective means to probe structure–property relationships.[57] In this work, RAC features were derived from MOF graphs using atomic identity (I), coordination environment and topology (T), Pauling electronegativity (χ), covalent radius (S), and nuclear charge (Z) as primary attributes, while polarizability (α) was additionally incorporated to describe linker-mediated interactions. The RAC formulation is given in eqn (1), and a total of 176 chemical descriptors were generated.

$$ {}^{\text{start}}_{\text{scope}}P_d^{\text{diff}} = \sum_i^{\text{start}} \sum_j^{\text{scope}} \left(P_i - P_j\right)\delta\left(d_{i,j} - d\right) \quad (1) $$

where ${}^{\text{start}}_{\text{scope}}P_d^{\text{diff}}$ denotes the aggregated difference in a given atomic property between atoms belonging to the start list and those in the corresponding scope list at a bond distance d. The atomic terms P_i and P_j correspond to the property values of atoms i (selected from the start atom list) and j (selected from the scope atom list), respectively, and $d_{i,j}$ defines their bond-based distance within the MOF graph. Molecular graph generation and the classification of start and scope atom sets were carried out according to the methodology reported by Moosavi *et al.*[56]

After feature extraction, 14 geometric descriptors and 176 chemical descriptors were obtained.

2.2 Adsorption capacity labels

The adsorption capacity of MOFs was determined using GCMC simulations. Six adsorption pressure conditions were established as labels based on typical methane adsorption applications such as methane storage, methane purification, and landfill gas treatment. Detailed adsorption conditions are summarized in Table 1. GCMC simulations were carried out with RASPA2 (ref. 58) to quantify methane uptake and analyze adsorption behavior in MOFs under six operating conditions. Nonbonded interactions between methane and framework atoms were described using the Lennard-Jones potential. Lennard-Jones parameters for different atom types were derived using Lorentz–Berthelot mixing rules, with methane parameters adopted from the study by Stubbs *et al.*[59] Framework atoms were described using the UFF.[60] Electrostatic interactions were described using REPEAT-derived partial charges, which were applied exclusively during framework optimization rather than during the GCMC simulations. Van der Waals forces were evaluated *via* an atom-based summation scheme. Each methane adsorption simulation began with 5000 Monte Carlo equilibration cycles, followed by 5000 production cycles. The number of moves per Monte Carlo cycle followed the protocol of Burner *et al.*,[51] including translational, rotational, swap, and exchange moves. Methane uptake was then obtained by averaging the number of adsorbed molecules over the production cycles of GCMC simulations.

Table 1 Adsorption conditions corresponding to six adsorption pressures used for model development

Task	Pressure	Temperature	Application
T1	0.01 bar	338 K	Landfill gas treatment
T2	0.9 bar	298 K	Methane purification
T3	4.4 bar	338 K	Landfill gas treatment
T4	5.8 bar	298 K	Methane separation and purification
T5	9 bar	298 K	Methane purification
T6	65 bar	298 K	Methane storage

2.3 Model architecture

MTL trains several related tasks within a unified framework, allowing knowledge to be shared among tasks during model optimization.[46] An encoder-focused architecture[61] with hard parameter sharing was adopted in this study. During feature extraction, identical network parameters were employed for all tasks to learn shared feature representations, while separate output layers were assigned to each task during decoding,[62] enabling task-specific predictions. Because adsorption behaviors at different pressures are intrinsically correlated, the six pressure specific tasks were jointly learned using a hard parameter sharing strategy to improve the training efficiency. To strengthen interactions between tasks and shared features, an attention mechanism module was embedded into the neural network, enhancing the weight of important features during training. The attention layer dynamically adjusted feature weights per task, enabling more sophisticated fusion of shared features and facilitating flexible handling of multiple related tasks. The model architecture is illustrated in Fig. 1, and more details can be found in the SI.

2.4 Working capacity screening mechanism

In industrial methane applications, PSA is commonly employed, involving adsorption under high pressure and desorption under low pressure to enable methane storage, transportation, and purification. Material performance evaluation depends not only on adsorption capacity at high pressure but also critically on the difference in adsorption capacity

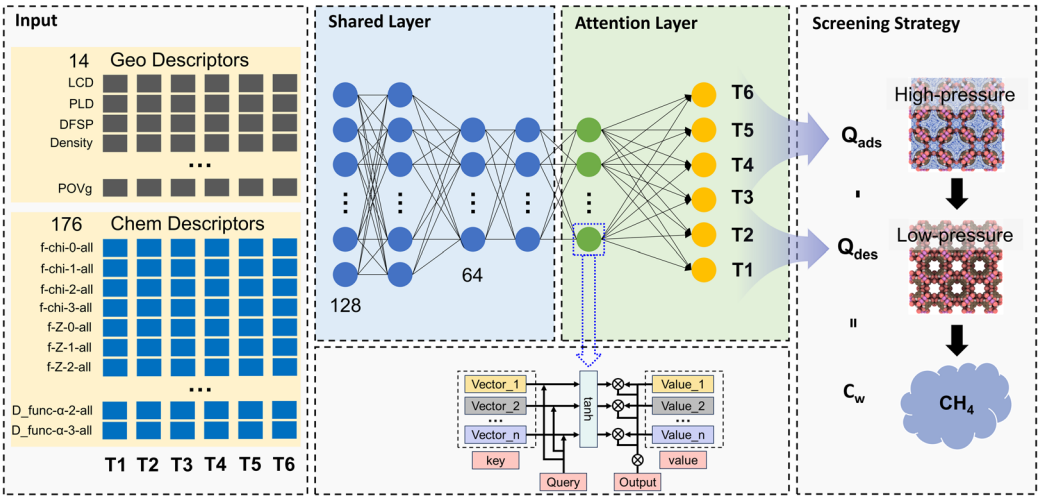

Fig. 1 Schematic diagram of the MTL model incorporating the attention mechanism and the PSA screening mechanism developed in this study.

between high and low pressures. Focusing solely on high-pressure adsorption without considering low-pressure desorption performance fails to accurately represent the material's practical value in cyclic operations. The proposed MTL model was integrated with a PSA screening mechanism. Based on the predicted adsorption capacities under high and low pressures, the working capacity of materials was calculated and employed as the key metric for identifying efficient and practical MOFs. The feature-sharing screening pattern of MTL under different pressure conditions is utilized to overcome the limitations of conventional single-pressure prediction methods. The working capacity calculation method is presented in eqn (2):

$$C_w = Q_{ads} - Q_{des} \quad (2)$$

where C_w denotes the working capacity, defined as gravimetric working capacity when expressed in g g^{-1} and as volumetric working capacity when expressed in cm^3 cm^{-3}; Q_{ads} represents the adsorption amount at a higher operational pressure during industrial adsorption processes; Q_{des} corresponds to the adsorption amount at a lower pressure during desorption.

2.5. Model training and performance evaluation

Methane adsorption capacities of MOFs measured under six pressure conditions were used to assemble the training dataset. Geometric and chemical descriptors served as input features, while methane adsorption capacities under the six pressure conditions were used as response labels. The dataset was randomly divided into free training and test subsets at a ratio of 8 : 2. Model robustness was evaluated using five-fold cross-validation on the free training set, with each fold optimized over 100 training epochs.

A series of optimization strategies were implemented to enhance the performance of the MTL model. The adaptive learning rate optimizer Adam was employed to adjust the learning rate, with the mean squared error (MSE) selected as the loss function. To address varying task importance in MTL, a dynamic task weighting mechanism was applied, in which task weights were updated inversely proportional to their losses. Tasks exhibiting higher losses received increased weights, whereas those with lower losses were assigned reduced weights. A total of 200 training epochs were set to ensure sufficient iteration and improve model generalization. Adjustable batch sizes of 32, 64, and 128 were tested to control the number of samples used per parameter update. Dropout layers were incorporated into the shared layers to prevent overfitting by randomly deactivating a proportion of neurons, thereby enhancing model robustness. Dropout rates ranging from 0.2 to 0.7 were determined *via* hyperparameter optimization.

Model performance was evaluated using the coefficient of determination (R^2) and root mean squared error (RMSE). R^2 quantifies the proportion of variance in the data explained by the model, with higher values indicating a better fit to the observed data. The RMSE reflects the precision of model predictions and places greater emphasis on larger deviations, providing insight into the model's performance on extreme values. The formulae of R^2 and RMSE are provided in eqn (3) and (4), respectively.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \quad (3)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \quad (4)$$

where y_i represents the true value, $\hat{y}_i$ denotes the predicted value, n is the number of samples, and $\bar{y}$ is the mean of the true values.

The resulting model enables the prediction of methane uptake and working capacity, guiding the identification of MOFs with superior performance.

2.6 Model interpretation and optimization

The contributions of descriptors were interpreted by SHapley Additive exPlanations (SHAP).[63] The overall average importance

of each descriptor on the model outcomes was obtained by calculating and averaging the Shapley values of all data points. Analysis of both global and local feature importance, combined with SHAP dependence analysis, facilitated the identification of potential avenues for model improvement, such as recognizing unimportant or redundant features to simplify the model or enhance its performance. The contributions of descriptors to the prediction were ranked, and the optimal descriptor ranges for enhancing the methane adsorption capacity in MOFs were identified, aiding the design of high-performance materials. The calculation method for Shapley values is presented in eqn (5).

$$\varnothing_i = \sum_{S \subseteq n/i} \frac{|S|!(n - |S| - 1)!}{n!} [f(S \cup i) - f(S)] \quad (5)$$

where $\varnothing_i$ denotes the Shapley value predicted for feature i; n represents the complete feature set; S refers to a feature subset that excludes i; and $f(S)$ indicates the corresponding model prediction evaluated on S.

Descriptor selection in the initial model development often included redundant features. After calculating descriptor importance, the contribution rankings for the six training tasks were established. A comprehensive importance ranking was obtained by averaging the absolute importance values of each descriptor across the six tasks. Based on the integrated importance ranking, the top twelve descriptors were retained and used to rebuild the prediction model. Model retraining was performed on the original dataset with the same hyperparameter optimization scheme, aiming to streamline the feature space, decrease computational demand, and improve generalization performance.

3. Results and discussion

3.1 Descriptive statistical analysis

As shown in Fig. 2, scatter plots and boxplots summarize 14 geometric descriptors, 5 representative chemical descriptors, and methane adsorption capacities of MOF samples under six different pressures within the dataset. Fig. 2A–C reveal pronounced right-skewed and long-tailed distributions for these chemical descriptors. The atom identity descriptors f-lig-I-1 and f-lig-I-2 exhibit mean values of 69.90 and 112.50, respectively, with some samples reaching maximum values of 1210 and 1920, indicating considerable variability. Topology descriptors mc-T-1-all and f-T-3-all display wide ranges from 8 to 739.14 and up to 38 976, respectively, while their means concentrate at lower values of 47.83 and 483.42. Although most samples exhibit coordination environments in the lower range, the long tails indicate a minority with highly complex coordination networks corresponding to functionally unique MOFs. The nuclear charge descriptor f-lig-Z-1 spans from 78.86 to 33 248, with a mean of 2190.62, reflecting significant diversity in ligand elemental composition across a wide range of low to high nuclear charges. Pauling electronegativity descriptors f-lig-chi-3 and f-lig-chi-0 show mean values of 945.68 and 232.41, with maxima reaching 15 631.39 and 3332.93, respectively. Covalent radius descriptors F-lig-S-2 and F-lig-S-3 present mean values of 53.58 and 64.42, with maximum values of 920.44 and 1121.64, respectively. Overall, the numerical distributions of these descriptors reflect the high diversity of MOF ligands in atomic composition, topology, and nuclear properties. This diversity may offer valuable reference points for further understanding

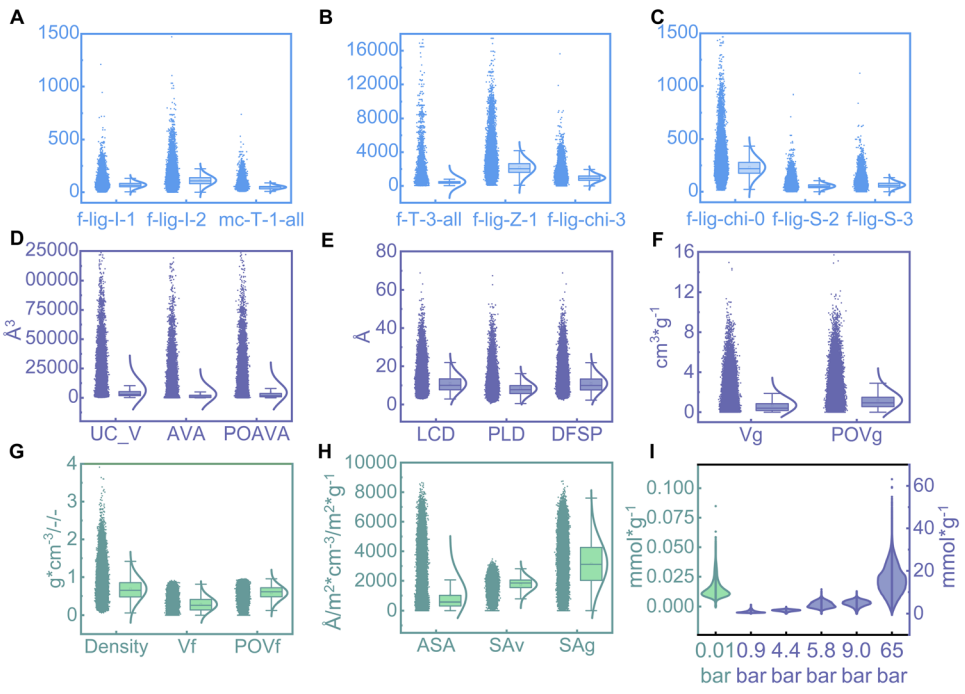

Fig. 2 Boxplots with normal distribution curves of descriptors and methane adsorption labels in the dataset: (A)–(C) chemical descriptors; (D)–(H) geometric descriptors; (I) methane adsorption capacities at six pressures.

structure–performance relationships and for guiding the screening of high-performance MOFs.

Statistical analysis of UCv and related volumetric descriptors for MOF samples revealed wide distribution ranges for UCv, AVA, and POAVA, spanning from 0.4407 Å^{3}to 198 068 Å^3, as illustrated in Fig. 2D. The means of these descriptors concentrate in the lower range, measured at 6310.63 Å^3, 3007.67 Å^3, and 4507.69 Å^3, respectively, with a small number of MOF samples significantly exceeding these averages. The UCv and related volumetric descriptors display clear right-skewed distributions, indicating that most MOF samples cluster within smaller UCv, while a minority with exceptionally large volumes elevate the maximum values substantially. Similar distribution patterns appear in pore size descriptors LCD, PLD, and DFSP, as shown in Fig. 2E, with values ranging from approximately 0.50 to 68.87 Å. This wide span reflects the diversity of pore structures within the sample set, covering very small micropores to larger channels. The means of these descriptors are 11.05 Å, 8.33 Å, and 10.98 Å, respectively, with distributions also right-skewed. The majority of pore sizes cluster at smaller scales, while a few samples exhibit notably larger pores. Distributions of volumetric descriptors Vg and POVg, shown in Fig. 2F, range from approximately 0.00025 to 15.70 cm^3 g^{-1}, demonstrating significant variability in pore volumes among the samples. Mean values register at 0.68 cm^3 g^{-1} and 1.19 cm^3 g^{-1}, with overall right-skewed distributions. Fig. 2G and H present distributions for other six descriptors. SAv ranges from 42.93 to 3474.6 m^2 cm^{-3}, and SAg ranges from 23.33 to 8748.9 m^2 g^{-1}. Both exhibit wide distribution ranges with means of 1745.01 m^2 cm^{-3} and 3153.68 m^2 g^{-1}, respectively, reflecting considerable diversity in pore surface areas across the dataset.

Fig. 2I illustrates the distribution of methane adsorption capacities for MOFs under six pressure conditions. Overall, the methane adsorption capacity shows a clear increasing trend with rising pressure, as both the mean and median values increase progressively. The median adsorption capacity increases from approximately 0.013 mmol g^{-1} under low pressure to about 15.14 mmol g^{-1} at high pressure, indicating a significant enhancing effect of pressure on adsorption performance. The distribution under low pressures (0.01 bar to 9 bar) is relatively concentrated, as evidenced by narrow and sharply peaked violin plots. In contrast, at 65 bar, the adsorption capacity distribution becomes more dispersed, reflecting greater diversity among MOF samples and a broader range of adsorption capability distribution under high-pressure conditions.[64] At the highest pressure of 65 bar, a pronounced long-tail distribution appears, where several samples exhibit adsorption capacities far exceeding the majority. These outliers correspond to MOFs with unique structures or exceptional pore characteristics, which play a critical role in elevating overall performance. The distributions of methane adsorption capacity across varying pressures not only demonstrate the promoting effect of pressure but also reveal performance differences and diversity among samples. These insights may provide a strong foundation for a deeper understanding of MOF adsorption behavior and for optimizing screening strategies.

Table 2 Model prediction performance of methane adsorption

	Full-feature model		Feature-optimized model		Feature values	
Task	R^2	RMSE	R^2	RMSE	Min value	Max value
T1	0.902	0.002	0.816	0.003	1.0×10^{-6}	0.085
T2	0.924	0.157	0.856	0.216	3.0×10^{-6}	4.197
T3	0.920	0.127	0.866	0.164	2.0×10^{-5}	3.204
T4	0.924	0.398	0.852	0.554	9.4×10^{-5}	11.302
T5	0.941	0.349	0.895	0.467	7.9×10^{-5}	10.384
T6	0.989	0.716	0.984	0.848	4.0×10^{-4}	63.146

3.2 Prediction performance of the full-feature model for single pressure adsorption

The test set comprising 50 470 samples excluded from model training was employed to evaluate the predictive performance of the MTL model. The training loss curve of the model is shown in Fig. S1A. Table 2 summarizes the test results of the full-feature model and the optimized model with simplified features across six adsorption prediction tasks. The model achieved R^2 values exceeding 0.900 on all six tasks, indicating strong data fitting across all prediction targets. The highest R^2 of 0.989 was obtained for T6, with R^2 values generally increasing as the pressure associated with the prediction tasks increased. RMSE values should be interpreted in the context of each task's label range. Overall, RMSE results across all six tasks remain within acceptable limits, reflecting low prediction errors. Fig. 3 compares the true values and predicted methane adsorption capacities across the six tasks of the MTL model. Data points for T6 predominantly align along the diagonal, demonstrating strong predictive consistency for this high-pressure task. Other tasks exhibit varying degrees of outliers, notably under low-pressure adsorption conditions (*e.g.*, T1, T2, T3, and T4), indicating reduced predictive consistency compared to high-

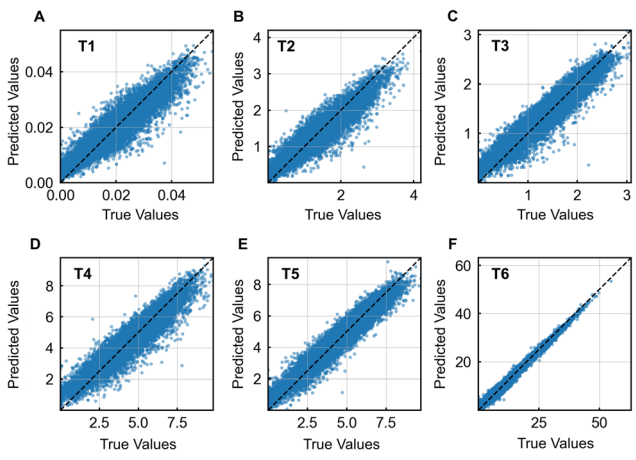

Fig. 3 Distribution of predicted and true values for the full-feature model: (A) T1; (B) T2; (C) T3; (D) T4; (E) T5; (F) T6.

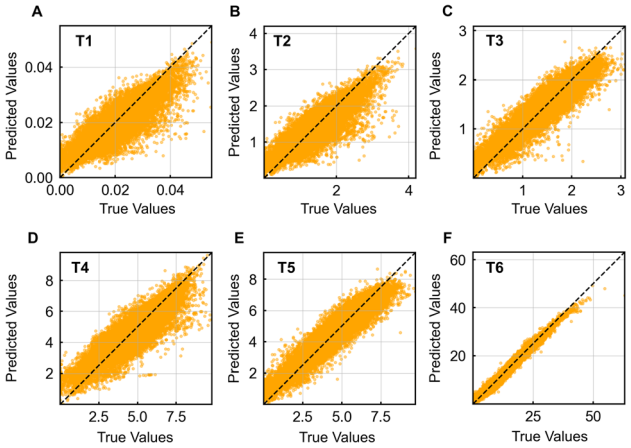

Fig. 4 Distribution of predicted and true values for the feature-optimized model: (A) T1; (B) T2; (C) T3; (D) T4; (E) T5; (F) T6.

pressure tasks. Fig. S2 presents the R^2 training curves, which rise rapidly during early epochs, reflecting rapid feature learning by the model. The R^2 curves stabilize with increasing epochs, exhibiting minimal fluctuations. Consistent trends between training and test sets suggest high fitting stability without signs of overfitting.

3.3 Prediction performance of the feature-optimized model for single pressure adsorption

By calculating the average absolute SHAP values across six tasks in the MTL model, the 12 most influential features were identified: Vf, POVf, f-lig-chi-0, f-lig-S-2, f-lig-I-2, density, SAg, SAv, f-lig-S-3, f-lig-chi-3, f-lig-Z-1, and f-lig-I-1. The prediction model was retrained using these top features. The comparison between predicted and true values for the feature-optimized model across six tasks is shown in Fig. 4. The training loss curve of the feature-optimized model is shown in Fig. S1B. Model prediction performance is summarized in Table 2, where R^2 values above 0.800 were achieved across all six tasks. The highest R^2 of 0.984 was observed for T6. Compared to the full-feature model, R^2 values of the feature-optimized model decreased by 0.046 to 0.086 for T1 through T5. Nevertheless, T6 maintained a high R^2 of 0.984. Similarly, the R^2 values of the feature-optimized model exhibit an increasing trend with the rise of pressure conditions associated with the prediction tasks, consistent with the pattern observed in the full-feature model. The feature-optimized model exhibited slightly increased RMSE values compared to the full-feature model, with increments ranging from 0.001 to 0.156. Overall, the simplified model showed reduced R^2, especially under low-pressure conditions, although the lowest R^2 of 0.816 for T1 still indicates acceptable performance. The RMSE results corroborate a modest increase in prediction errors for the feature simplified model. Fig. 5 illustrates the comparison between true values and predicted values of methane adsorption capacities across six tasks using the feature-optimized model. Similar to the full-feature model, data points for T6 mostly cluster around the diagonal line with few outliers, indicating strong predictive consistency for this task. In contrast, T1 to T5 exhibit noticeably more outliers, predominantly located in the lower-right region, indicating a tendency of the model to underestimate the adsorption performance of certain MOFs. Fig. S3 shows the R^2 training curves for the feature-optimized model. The curves for both training and test sets align closely, rising rapidly before stabilizing, which indicates high fitting stability.

3.4 Working capacity screening performance

The MTL screening framework proposed in this study offers an improved approach for selecting materials suitable for engineering applications. A screening process was performed on a new dataset, consisting of 50 470 MOF samples, which were not included in the model training. Working capacities were calculated at two PSA pressure ranges: PSA1 (65 bar to 5.8 bar) and PSA 2 (9 bar to 0.9 bar). The performance of models are summarized in Table 3. Both the full-feature and feature-optimized models achieved R^2 values exceeding 0.9. The highest R^2 of 0.992 was obtained by the full-feature model for PSA1 screening.

3.5 Comprehensive feature importance

To clarify the contribution of each descriptor to the prediction results and identify descriptors with significant impact on

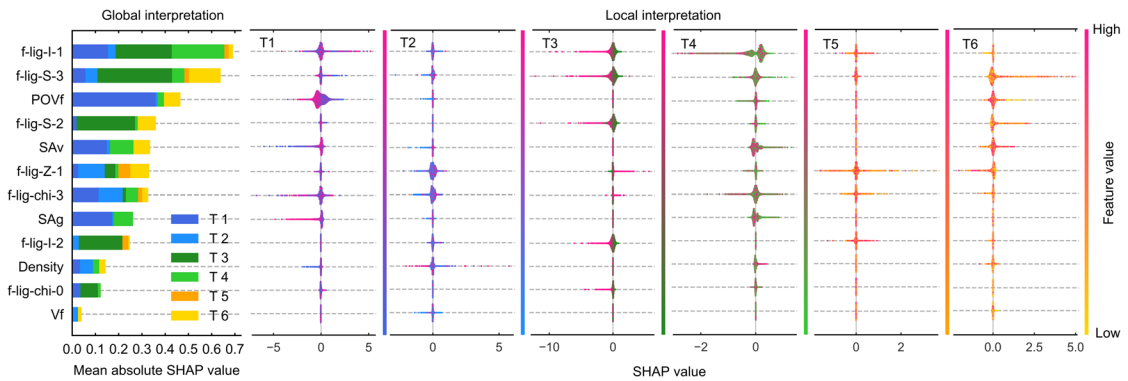

Fig. 5 Comprehensive global interpretation (average feature importance) and comprehensive local interpretation (SHAP value distribution) for the six prediction tasks.

Table 3 Model performance in predicting the working capacity

		PSA1		PSA2	
Working capacity	Models	R^2	RMSE	R^2	RMSE
Gravimetric working capacity	Full-feature model	0.992	0.622	0.953	0.320
	Feature-optimized model	0.991	0.693	0.924	0.407
Volumetric working capacity	Full-feature model	0.962	8.724	0.931	5.448
	Feature-optimized model	0.951	9.891	0.894	6.725

methane adsorption prediction, SHAP analysis was applied to the methane adsorption prediction model. Fig. 5 presents the top twelve features ranked by the average absolute SHAP values across six tasks, alongside SHAP summary plots for each individual task. These illustrate the descriptors with the highest overall importance and their influence on model predictions. The descriptor f-lig-I-1 showed the greatest importance, followed by f-lig-S-3, POVf, f-lig-S-2, and SAv. Both chemical and geometric descriptors contributed significantly to adsorption performance. Among chemical descriptors, functional group-related descriptors (f-lig-) ranked prominently, indicating that the local chemical environment around functional groups has a significant impact on methane adsorption. This highlights the role of active sites and chemical properties in regulating adsorption performance.[65] Geometric descriptors such as POVf, SAv, SAg, density, and Vf also ranked within the top twelve, demonstrating that the pore structure and surface area exert considerable influence on methane adsorption. This reflects the combined effects of chemical and structural characteristics on MOFs' adsorption performance.

The descriptor f-lig-I-1 has the highest rank in overall feature importance. It characterizes the chemical environment of the first coordination shell surrounding the core atoms of MOF functional groups.[55] The atomic composition of this shell directly reflects the local changes at adsorption sites, thereby influencing their activity. Higher f-lig-I-1 values indicate a greater number of bonds within the ligand and a larger ligand size, which impacts the pore size and shape near adsorption sites. Local importance plots reveal that f-lig-I-1 exerts both positive and negative effects, with a generally positive influence on predicted adsorption capacity in T4. f-lig-S-3 and f-lig-S-2 ranked second and fourth in overall importance, respectively. These represent statistical measures of covalent radii for atoms located in the third coordination shell and beyond, and the second coordination shell, around the functional group's core atoms. They capture the volume and spatial environment of atoms at middle and distal distances, affecting MOFs' pore shape and size, which subsequently influence methane accessibility and diffusion. These findings suggest that f-lig-I-1 influences the electronic properties of adsorption sites in MOF samples for methane adsorption, while f-lig-S-3 and f-lig-S-2 represent covalent radii reflecting the spatial distribution of ligand atom sizes within the molecular topology. These descriptors affect the spatial framework of MOFs, determining the molecular accessibility and binding efficiency. The high importance of these three chemical descriptors indicates that capturing key factors governing adsorption performance requires consideration of both the local electronic environment near functional groups and the mid-to-long-range spatial structure.

Among geometric descriptors, POVf and SAv ranked third and fifth, respectively. POVf denotes the fraction of pore-accessible volume, representing the effective pore space available for methane molecules.[66] SAv corresponds to the surface area per unit volume, which directly relates to the number of available adsorption sites.[67] Other geometric descriptors among the top twelve include SAg, density, and Vf. Although the model was trained using a substantially larger number of chemical descriptors than geometric ones (176 and 14), nearly half of the top twelve features were geometric descriptors (5 in total). This highlights the significant role of geometric features in prediction performance. The prominence of these geometric descriptors underscores the importance of effective pore volume, surface area, and density related metrics for methane adsorption in MOFs, consistent with conclusions drawn from extensive prior studies.[45,68] From the perspective of integrated multi-task descriptor importance, factors influencing methane adsorption performance range from the local chemical environment of functional groups at the microscopic scale, through mid- and long-distance spatial constraints, to macroscopic effective pore volume and surface area. These insights should guide the design of MOFs optimized for high-performance methane adsorption.

3.6 Analysis of multi-task consistency and differences

Fig. 6 presents the feature importance and SHAP summary plots for six predictive tasks under varying pressure conditions. Under low pressure (0.9 bar, Fig. 6B), chemical descriptors such as f-I-3-all, f-I-1-all, and lc-chi-3-all ranked prominently among the key features, indicating that the topological connectivity of functional group neighboring and distant atoms, as well as the electronegativity of linker atoms, gradually became key factors influencing methane adsorption performance in MOFs. Local importance plots showed these three descriptor types exerted both positive and negative effects on adsorption performance.

As pressure increased to intermediate levels (4.4 bar, Fig. 6C; 5.8 bar, Fig. 6D; 9 bar, Fig. 6E), the MTL model revealed consistent patterns in feature importance rankings for predicting the methane adsorption capacity. Topological and chemical differential descriptors associated with various atom types, metal-centered, functional group, and linker-centered, such as D_mc-T-3-all, D_func-l-3-all, and D_lc-Z-3-all, along

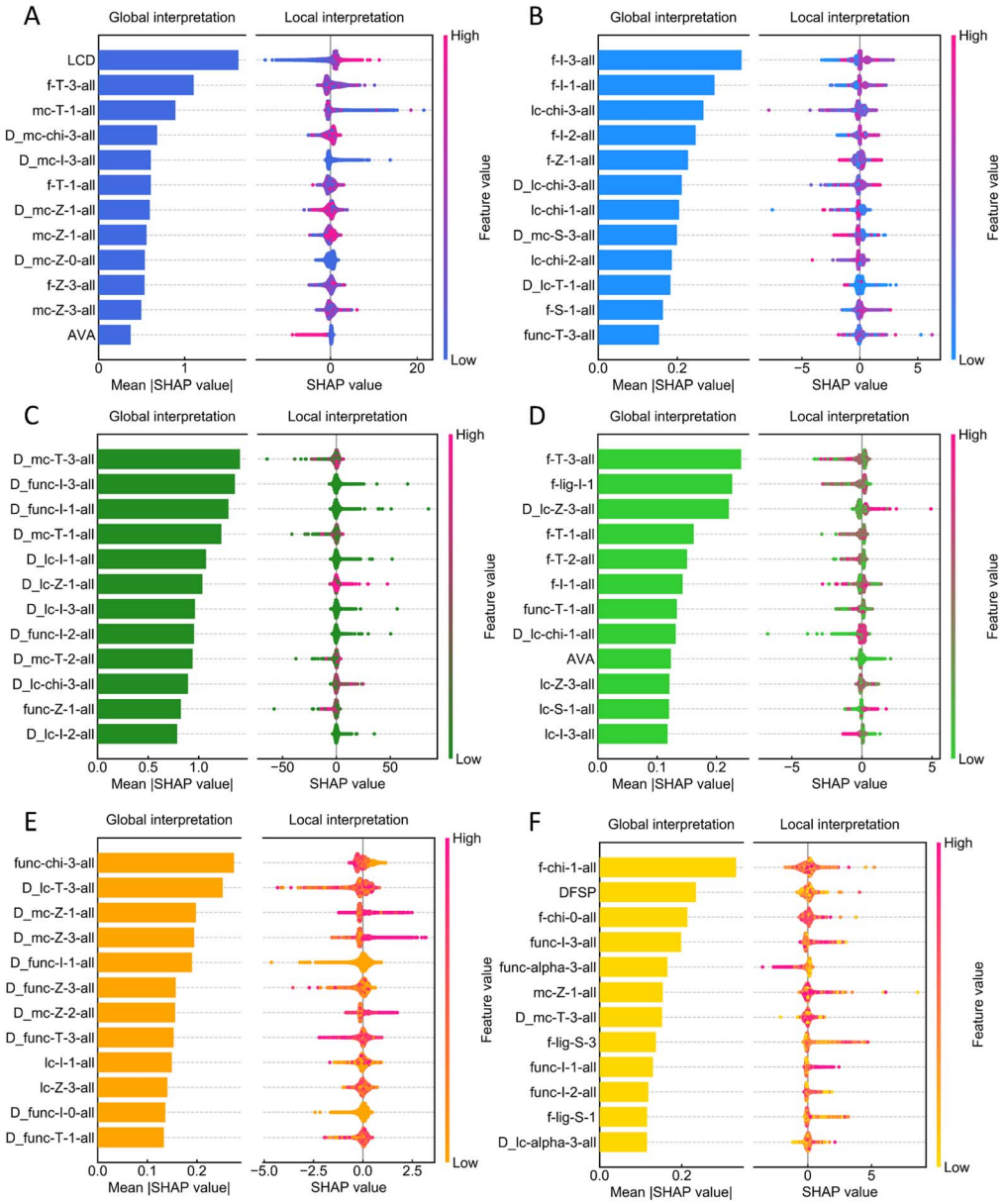

Fig. 6 Global interpretation and local interpretation for the six predictive tasks: (A) T1; (B) T2; (C) T3; (D) T4; (E) T5; (F) T6.

with descriptors capturing medium- and long-range topological and chemical environments (f-T-3-all, f-lig-I-1, and func-chi-3-all), dominated the rankings. These differential features indicate that methane adsorption performance at intermediate pressures strongly depends on local chemical heterogeneity and the spatial distribution of electronic structures within the material.

Notably, the relative importance of geometric descriptors declined with increasing pressure, replaced by multilayered chemical descriptors grounded in topological differentials and electronic properties. This finding underscores the importance of chemical modification of functional groups and multiscale optimization of linker electronic structures for enhancing methane adsorption at elevated pressures. The top-ranked descriptors were f-chi-1-all, DFSP, and f-chi-0-all in T6 (Fig. 6F), with normalized mean absolute SHAP values of 0.33, 0.23, and 0.21, respectively. This result highlights the dominant role of functional group electronegativity under high pressure adsorption, reflecting the modulation of local electronic environments that influence the polarization ability of adsorption sites and thus affect the interaction strength with methane molecules.[69] Under high pressure adsorption conditions, the geometric descriptor DFSP also maintained high importance, indicating that pore connectivity and methane accessibility remain critical factors even at near-saturation adsorption stages. Effective enhancement of methane adsorption performance under high pressure requires concurrent optimization of functional group electronic environments and reasonable pore structure design to ensure smooth methane ingress and efficient filling.[70]

3.7 Dependency analysis

Potential interactions among descriptors were identified through comparison of feature SHAP values. SHAP dependence under feature interactions was analyzed for the top-ranked descriptors. Fig. S4 to S9 summarize the SHAP dependence plots, illustrating interactions among the six highest-ranking descriptors across six predictive tasks. Analysis of overall SHAP dependencies indicated that geometric descriptors typically exhibited clear positive or negative contributions, whereas chemical descriptors displayed more complex patterns, contributing both positively and negatively to adsorption performance.

Further investigation focused on SHAP dependence relationships for descriptors suspected to exhibit feature interactions based on prior importance analyses. Fig. 7 illustrates the SHAP dependence relationships capturing the interaction between the descriptor f-lig-I-1 and geometric descriptors related to pore size and shape in T3. Similar interaction patterns were observed between f-lig-I-1 and the three pore-related geometric descriptors. When geometric descriptor values were small, the positive contribution of f-lig-I-1 to methane adsorption gradually diminished as its value increased. After f-lig-I-1 exceeds 32, SHAP values shifted to negative, and the negative contribution to the adsorption performance gradually increases. When geometric descriptor values were high, f-lig-I-1 consistently exhibited negative contributions, with the magnitude of negative impact increasing more moderately as the descriptor value increased.

3.8 Feature synergy analysis

SHAP dependence analysis revealed that geometric descriptors exhibit more pronounced interaction effects compared to chemical descriptors. To further investigate interactions among descriptors and their combined influence on methane adsorption capacity, interactions among the top-ranked descriptors Vf, POVf, SAv, SAg, and density were analyzed. Fig. 8A and B present the synergistic effects of POVf, SAv, and density on methane adsorption in MOFs for the T3 (4.4 bar) and T6 (65 bar) tasks, respectively. The results indicate that, in the T3 task, samples with high adsorption capacity were mainly concentrated within specific density ranges. The impact of SAv on adsorption was more complex, with high-performance MOFs observed at both high and low SAv values. At low density (<1 g cm^{-3}), the adsorption capacity increased gradually with increasing POVf.

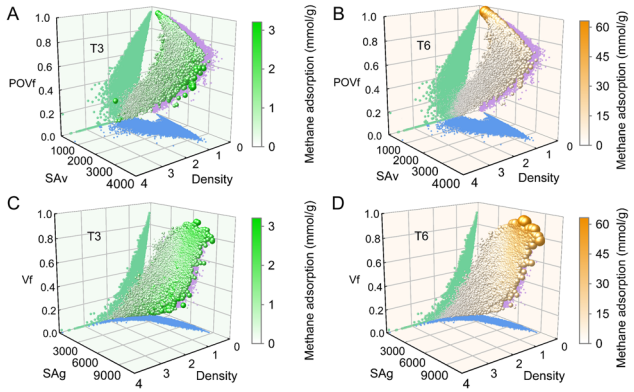

Fig. 8 Three-dimensional feature interaction analysis for methane adsorption. (A and B) The interaction effects of POVf, SAv, and density for tasks T3 and T6; (C and D) The interaction effects of Vf, SAg, and density for tasks T3 and T6. Methane adsorption capacity is encoded by both the marker size and color intensity in the 3D scatter plots.

At higher densities (>1 g cm^{-3}), adsorption initially increased sharply with increasing POVf, then gradually declined. For task T6, MOFs with high adsorption were mainly found under low density and high POVf conditions. At lower density, methane uptake showed opposite trends with respect to SAv and POVf, decreasing as SAv increased while exhibiting a positive dependence on POVf. At high density, adsorption also increased gradually with increasing SAv. These observations indicate that methane adsorption in MOFs is optimal for specific ranges of POVf, density, and SAv. This highlights the significant regulatory roles these structural parameters have on adsorption capacity.

Fig. 8C and D depict the synergistic effects of Vf, SAg, and density on methane adsorption in MOFs for the T3 (4.4 bar) and T6 (65 bar) tasks, respectively. The results show that, in the T3 task, high adsorption samples were mainly concentrated at low density (<1 g cm^{-3}) and within specific ranges of SAg and Vf. At high density (>1 g cm^{-3}), the adsorption capacity increased with increasing POVf. At low density, adsorption initially decreased and then increased as POVf increased. For the T6 task, samples with high adsorption were primarily found under conditions of low density and high SAg and Vf, with adsorption capacity increasing as the density decreased and SAg and Vf increased. Fig. S10 and S11 illustrate the interactive effects of four descriptors on methane adsorption for other tasks. The results indicate that, under high-pressure conditions, methane

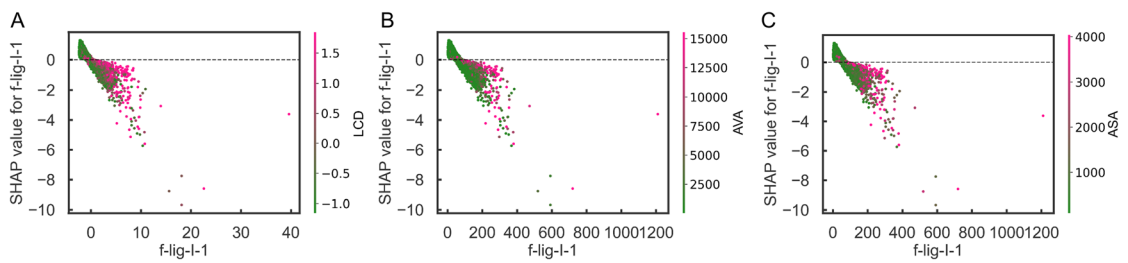

Fig. 7 SHAP dependence plot of T3 for f-lig-I-1 with interaction of (A) LCD; (B)AVA; (C) ASA.

adsorption in MOFs increased significantly with decreasing density and increasing POVf and Vf, exhibiting a clear linear correlation. In contrast, during low-pressure adsorption, the relationship between adsorption capacity and density, POVf, and Vf did not follow a simple monotonic pattern, but rather clustered within specific ranges of geometric descriptors, exhibiting nonlinear response characteristics. These findings indicate significant differences in the mechanisms by which geometric descriptors influence methane adsorption performance under different pressure conditions.

3.9 Feature-performance relationship

Analysis of the features elucidated the relationship between MOF structural attributes and both methane adsorption and working capacity, revealing clusters of top-performing MOFs within specific structural ranges. To elucidate feature–performance relationships, the influence of key descriptors on MOF performance was systematically analyzed. Fig. 9 illustrates the relationships between working capacities, methane uptake capacity and feature descriptors of MOFs. The working capacity of MOFs correlated positively with POVf, Vf, and SAg and negatively with density. A similar trend was observed for methane adsorption in the T6 task. The relationship between SAv and working capacity was nonlinear, with large variations in methane uptake observed within narrow SAv ranges. For example, within the SAv range of 397 to 399 m^2 cm^{-3}, MOF samples demonstrated working capacities of 60.14 and 0.97 mmol g^{-1}, representing a difference of 59.17 mmol g^{-1}. Data points were highly scattered, showing no clear correlation with adsorption capacity. In the T3 task, relationships between geometric descriptors and adsorption capacity were nonlinear, with distinct dispersion observed across feature levels. POVf and Vf exhibited a trend of initial increase, followed by decrease, and then another increase. Both SAv and density showed volcanic relationships with optimal adsorption performance within specific ranges.

Consistent trends were observed between performance and the atomic identity descriptor (f-lig-I-1), covalent radius descriptor (f-lig-S-3), and nuclear charge descriptor (f-lig-Z-1) in both working capacity and methane adsorption. Although high adsorption MOFs exhibited dispersed distributions across these descriptor attributes, clustering of high-performance MOFs was still detected within specific attribute ranges, primarily located at lower attribute values.

3.10 Screening of materials with optimal working capacity

Tables S1 and S2 list representative geometric descriptors, chemical descriptors, and topology features for the top ten MOFs by gravimetric working capacity under PSA1 and PSA2 conditions, respectively. Fig. 10 and S12 present crystal structures of these MOFs, exhibiting high gravimetric working capacities. Descriptor analysis of these MOFs indicated that the top PSA1 materials share similar geometric characteristics, particularly density, SAg, POVf, and SAv, with coefficients of variation of 0.12 g cm^{-3}, 0.09 m^2 g^{-1}, 0.001, and 0.16 m^2 among the top ten samples, respectively. The geometric descriptor values of these materials fall within previously identified combined descriptor ranges associated with high adsorption capacity, including a density of 0.074 ± 0.009 g cm^{-3}, SAg of 6739.28 ± 588.04 m^2 g^{-1}, POVf of 0.007 ± 0.948, and SAv of 509.76 ± 80.19 m^2. However, chemical descriptors listed in Table S2 show coefficients of variation exceeding 0.4, indicating high dispersion. This suggests that MOFs with specific geometric descriptors are more likely to exhibit superior gravimetric methane adsorption working capacity. While top-performing MOFs of PSA1 display diverse topologies, investigation of PSA2 materials reveals that certain topologies more effectively enhance the gravimetric working capacity. Among

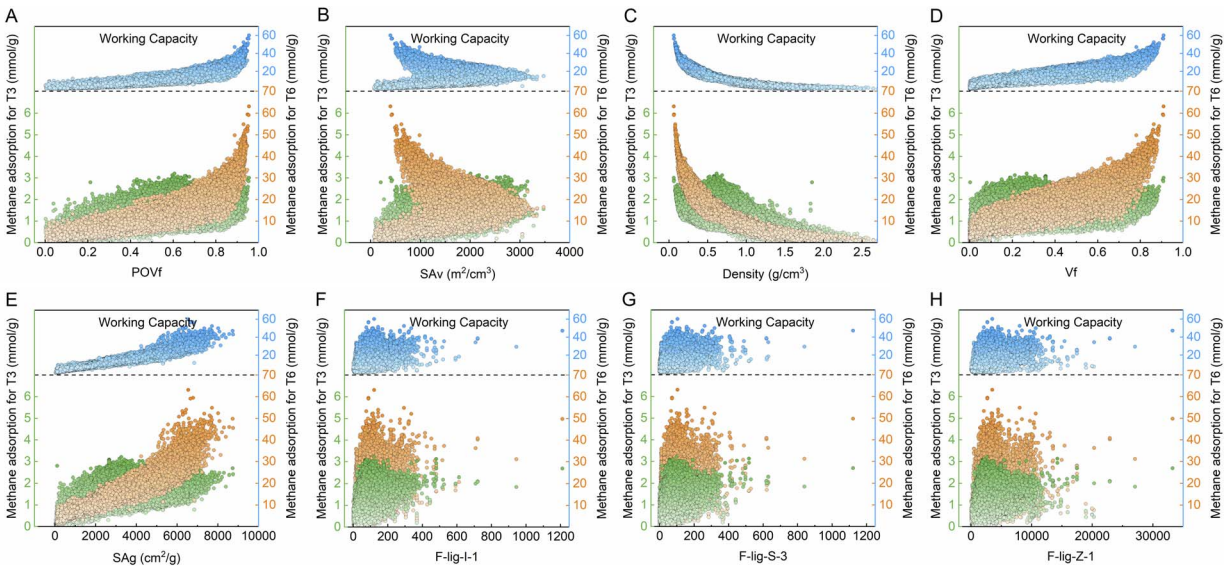

Fig. 9 The relationship between working capacity, methane adsorption capacity of MOFs and key structural features: (A) POVf; (B) SAv; (C) density; (D) Vf; (E) SAg; (F) f-lig-I-1; (G) f-lig-S-3; (H) f-lig-Z-1.

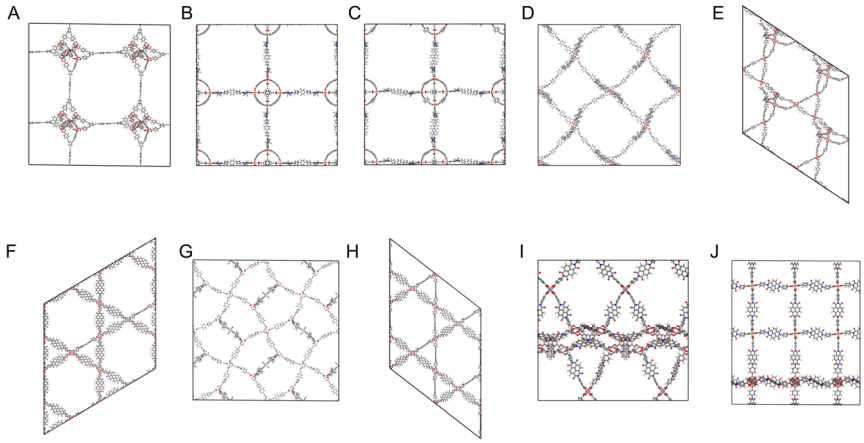

Fig. 10 Structures of top-performing MOFs for PSA1 by gravimetric working capacity: (A) m2_o33_o24_xaq; (B) ddmof_15865; (C) ddmof_16158; (D)ddmof_16861; (E)m2_o3_lcv; (F) m2_o29_nbo; (G) ddmof_21735; (H) m3_o29_nbo; (I) m3_o28_unh; (J) m2_o28_mou.

the ten structures listed in Table S3, eight possess the pcu topology, indicating a preference for pcu-type materials as efficient adsorbents for PSA2 gravimetric working capacity. Descriptor examination in Table S3 shows that PSA2 top-performing materials exhibit similar geometric descriptors, with coefficients of variation for density, SAg, POVf, and SAv at 0.08 g cm^{-3}, 0.11 m^2 g^{-1}, 0.09, and 0.13 m^2, respectively. Chemical descriptor variation coefficients all exceed 0.35, slightly higher than those observed for PSA1. Comparison of gravimetric working capacities reveals that the best value under PSA1 (0.929 g g^{-1}) substantially exceeds that of PSA2 (0.1438 g g^{-1}). Increasing the high-pressure segment in PSA adsorption significantly improves the methane working capacity.

Unlike automotive fuel tanks that focus on the gravimetric performance of adsorbents, industrial applications and methane storage prioritize the volumetric efficiency. Therefore, the volumetric working capacity serves as a crucial criterion for selecting high-performance materials. Tables S4 and S5 present the top-performing materials by volumetric working capacity for PSA1 and PSA2, along with their representative descriptors and topology features. Crystal structures of these high volumetric capacity MOFs are shown in Fig. S13 and S14. Comparison of volumetric working capacities indicates that the top value under PSA1 (213.45 cm^3 cm^{-3}) significantly exceeds that under PSA2 (113.07 cm^3 cm^{-3}). Descriptor analysis reveals significant differences between the optimal ranges of descriptors for top volumetric working capacity MOFs under PSA1 and those for top gravimetric working capacity MOFs. The most notable geometric descriptor ranges for volumetric capacity include a density of 0.48 ± 0.01 g cm^{-3}, SAg of 5193.60 ± 178.87 m^2 g^{-1}, POVf of 0.75 ± 0.01, and SAv of 2487.99 ± 113.93 m^2. This distinction highlights how gravimetric and volumetric criteria reflect different geometric structural features of high-capacity MOFs, where the former corresponds to properties of reduced mass per unit and the latter to properties of reduced volume per unit. Top-performing MOFs for volumetric working capacity exhibit similar characteristics, especially in geometric descriptors. The coefficients of variation for density, SAg, POVf, and SAv under PSA1 are 0.02 g cm^{-3}, 0.03 m^2 g^{-1}, 0.01, and 0.05 m^2, respectively, whereas those under PSA2 are 0.10 g cm^{-3}, 0.15 m^2 g^{-1}, 0.02, and 0.06 m^2, respectively.

Methane working capacities (desorption at 5 bar) of top-performing MOFs identified in this study were compared to those of experimentally synthesized nanoporous materials reported in the literature[71,72] under conditions of 65 bar and 298

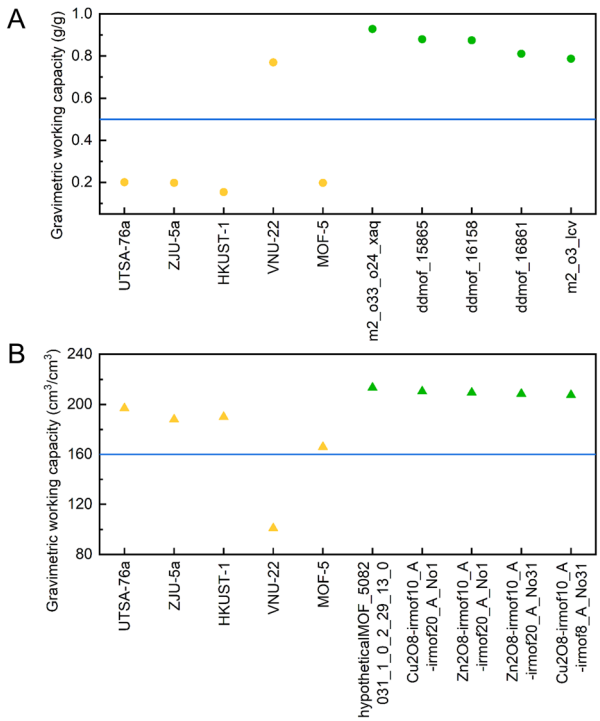

Fig. 11 Comparison of PSA1 working capacities between top-performing MOFs identified in this work and those reported in the published literature. (A) Gravimetric working capacity. (B) Volumetric working capacity. Yellow symbols represent representative high-performance materials reported in experimental studies, while green symbols correspond to MOFs with the highest working capacities identified in this study.

K.[73,74] The results are presented in Fig. 11. Fig. 11A shows that among the surveyed experimental MOFs, VNU-22 synthesized by Tu *et al.*[75] surpassed 0.5 g g^{-1} in gravimetric working capacity, exceeding the methane storage target established by the U.S. Department of Energy (DOE). The top-performing MOFs screened in this study demonstrated gravimetric working capacities exceeding that of VNU-22. Notably, m2_o33_o24_xaq exhibited the highest gravimetric working capacity at 0.929 g g^{-1}, with its crystal structure shown in Fig. 10A. Fig. 11B illustrates that UTSA-76a, ZJU-5a, and HKUST-1 all exceed 188 cm^3 cm^{-3} in volumetric working capacity, while hypotheticalMOF_5082031_1_0_2_29_13_0, identified in this study, achieved the highest volumetric working capacity of 213 cm^3 cm^{-3}. Its crystal structure is shown in Fig. S13A. These results demonstrate that the MOFs identified in this study possess the potential to surpass current synthesized materials in methane working capacity performance. Moreover, it should be noted that practical engineering applications also require consideration beyond adsorption capacity, including thermal and structural stability of materials.

Investigation of descriptors for materials with superior working capacity revealed that top-performing MOFs exhibit similar distributions in geometric descriptors such as density, SAg, POVf, and SAv. Based on the experimental findings, the following design guidelines are proposed to provide a valuable direction for future efforts aimed at developing MOFs with enhanced methane adsorption capacity and working capacity. (1) Target materials should possess SAv values in the range of 400 to 1500 m^2 cm^{-3} for optimal high-pressure adsorption performance. (2) Optimal performance is associated with SAg values between 5500 and 8700 cm^2 g^{-1} under high-pressure adsorption and between 1950 and 5700 cm^2 g^{-1} under low-pressure adsorption. Pursuit of maximum adsorption capacity necessitates minimizing density while maximizing Vf and POVf. (3) For gravimetric working capacity, MOFs with density around 0.074 g cm^{-3}, SAg near 6700 m^2 g^{-1}, POVf approximately 0.007, and SAv close to 500 m^2 demonstrate superior performance. MOFs optimized for volumetric working capacity exhibit values near density of 0.5 g cm^{-3}, SAg of 4700 m^2 g^{-1}, POVf of 0.75, and SAv of 0.40 m^2.

3.11 Applicability domain and limitations

The applicability of the proposed multi-task learning framework is defined by the chemical and structural space represented in the training database used in this study. The model was trained and validated on 252 352 chemically reasonable MOFs screened from eight publicly available databases. These MOFs cover a broad range of pore geometries, topologies, and chemical environments. However, predictions for materials that deviate substantially from this domain should be interpreted with caution, such as those possessing structural or chemical features not captured by the selected descriptors.

In addition to the geometric and chemical descriptor domain, the applicability of the model is constrained by the thermodynamic conditions under which the adsorption data were generated. All methane adsorption and working capacity data used for model training and validation were obtained at the temperatures considered in this study under six pressure conditions corresponding to three representative PSA application scenarios. Specifically, these conditions include: landfill gas treatment (4.4–0.01 bar, 338 K), methane purification (9–0.9 bar, 298 K), and methane storage, separation, and purification (65–5.8 bar, 298 K). Model predictions are therefore strictly applicable only to these defined pressure-temperature conditions.

It is necessary to acknowledge the limitations of the present study. Although several strategies were employed to mitigate potential overfitting, including five-fold cross-validation and performance evaluation using an independent test set obtained through random sampling that was not involved in model training (but originated from the same source as the training data), it should be noted that independent external validation using datasets outside the training database has not yet been conducted. Consequently, the current model is primarily validated within the chemical and structural space represented by the available MOF dataset.

Validation using independently sourced external datasets is essential for fully substantiating the capability of the proposed model to identify genuinely novel high-performance MOFs beyond its training domain. In future work, external validation will focus on newly reported or newly released experimental MOFs with available methane adsorption or working capacity data. First, compiling methane working capacity data for MOFs that are newly discovered or experimentally synthesized in subsequent studies to further validate and refine the proposed multi-task learning model. Second, the applicability of the model will be extended by incorporating working capacity prediction tasks under a broader range of pressure–temperature conditions, thereby enabling a more comprehensive assessment of model applicability and providing further insights into its generalization behaviour under extended operating conditions.

4. Conclusion

The design and screening of materials with superior methane storage capabilities remain critical challenges for advancing energy applications. This study presents the first integration of MTL with molecular simulation techniques to develop effective prediction models for methane adsorption performance and working capacity screening of MOFs. The MTL model capitalized on enhanced training efficiency and simultaneous multi-dimensional predictions, outperforming traditional machine learning methods and GCMC simulations in both accuracy and computational efficiency. The MTL-based methane adsorption capacity prediction model, trained on the dataset, achieved the highest R^2 of 0.989 across pressures of 0.01, 0.9, 4.4, 5.8, 9, and 65 bar. For working capacity screening under 65 bar adsorption and 5.8 bar desorption, full-feature and simplified models achieved R^2 values of 0.992 and 0.991, respectively, for gravimetric working capacity, while volumetric working capacity screening models reached R^2 values of 0.962 and 0.951, respectively. Subsequently, SHAP analysis was applied to assess

descriptor importance influencing methane adsorption of MOFs across different pressures within the MTL framework. Distinct differences in descriptor importance were observed between high-pressure and low-pressure conditions. Local importance and dependence analyses revealed strong correlations between certain geometric descriptors and adsorption capacity, with others exhibiting volcanic relationships. Moreover, quantitative analyses were conducted to elucidate the feature-performance relationships between adsorption capacity and descriptors. Among the top twelve features ranked by integrated importance, geometric descriptors accounted for 42%, with POVf, SAv, SAg, density, and Vf identified as the most critical geometric descriptors. The contribution of chemical descriptors proved more complex, as most showed significant dispersion relative to adsorption capacity. Only atomic identity (f-lig-I-1), covalent radius (f-lig-S-3), and nuclear charge (f-lig-Z-1) descriptors exhibited notable clustering of high-performance samples. Feature importance results indicate that pore connectivity and methane molecule accessibility, as captured by geometric descriptors, remain key determinants of methane adsorption, especially under high-pressure conditions. Based on these findings, the optimal feature range for developing high-capacity MOFs has been proposed. The results of this study contribute practical insights toward a comprehensive understanding of methane adsorption performance in MOFs. More importantly, these findings are expected to accelerate the development and identification of novel MOF adsorbents exhibiting exceptional gravimetric or volumetric working capacities across diverse practical applications.

Author contributions

Junhui Kou: methodology, software, writing – original draft. Tianle Liu: conceptualization, funding acquisition, writing – review & editing. Guosheng Jiang: funding acquisition, validation. Guokun Yang: validation, writing – review & editing. Zerang Li: investigation, methodology. Xiaoyang Ni: funding acquisition, resources, validation. Hao Liu: methodology, writing – review & editing.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The datasets and source code supporting this study can be found online at https://doi.org/10.5281/zenodo.18522489. Other data supporting this article have been included as part of the supplementary information (SI). Supplementary information: the definition of geometric descriptors; model architecture details; training loss curves and R2 training curve for the full-feature models and feature-optimized models; SHAP dependence plots of T1 to T6 for interactions among the top 6 important descriptors; 3D interaction diagram of features for T1, T2, T4, and T5; MOFs ranking of gravimetric and volumetric working capacities for PSA1 and PSA2; structures of top-performing MOFs by gravimetric and volumetric working capacity. See DOI: https://doi.org/10.1039/d5ta10538b.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2022YFC3004804).

References

1 S. Thakur and H. Solanki, *Int. Assoc. Biol. Comput. Dig.*, 2022, **1**, 275–281.
2 S. E. M. Fletcher and H. Schaefer, *Science*, 2019, **364**, 932–933.
3 M. Toghyani, S. Dadkhahfar and A. Alishahi, in. *Advances and technology development in greenhouse gases: emission, capture and conversion*, Elsevier, 2024, pp. 463–510.
4 H. Roohollahi, H. Zeinalzadeh and H. Kazemian, *Ind. Eng. Chem. Res.*, 2022, **61**, 10555–10586.
5 O. T. Qazvini, R. Babarao and S. G. Telfer, *Nat. Commun.*, 2021, **12**, 197.
6 I. Alali, A. U. Shehu and R. Mokaya, *Energy Environ. Sci.*, 2024, **17**, 5024–5038.
7 K. Nath, K. R. Wright, A. Ahmed, D. J. Siegel and A. J. Matzger, *J. Am. Chem. Soc.*, 2024, **146**, 10517–10523.
8 M. Ding, R. W. Flaig, H.-L. Jiang and O. M. Yaghi, *Chem. Soc. Rev.*, 2019, **48**, 2783–2828.
9 L. Tang, Q. Xu, Y. Zhang, W. Chen and M. Wu, *Electrochem. Energy Rev.*, 2022, **5**, 32–81.
10 F. Zhang, J. Zhang, B. Zhang, L. Zheng, X. Cheng, Q. Wan, B. Han and J. Zhang, *Nat. Commun.*, 2020, **11**, 1431.
11 W. Xie, Q. Fu, L. Yang, L. Yan, J. Zhang and X. Zhao, *ChemSusChem*, 2025, **18**, e202401382.
12 J. Wang, J. Liu, H. Wang, M. Zhou, G. Ke, L. Zhang, J. Wu, Z. Gao and D. Lu, *Nat. Commun.*, 2024, **15**, 1904.
13 P. G. Boyd, A. Chidambaram, E. García-Díez, C. P. Ireland, T. D. Daff, R. Bounds, A. Gładysiak, P. Schouwink, S. M. Moosavi, M. M. Maroto-Valer, J. A. Reimer, J. A. R. Navarro, T. K. Woo, S. Garcia, K. C. Stylianou and B. Smit, *Nature*, 2019, **576**, 253–256.
14 D. Li, L. Chen, G. Liu, Z. Yuan, B. Li, X. Zhang and J. Wei, *New Carbon Mater.*, 2021, **36**, 468–496.
15 A. Uzun and S. Keskin, *Prog. Surf. Sci.*, 2014, **89**, 56–79.
16 M. Pardakhti, P. Nanda and R. Srivastava, *J. Phys. Chem. C*, 2020, **124**, 4534–4544.
17 W. Gong, Y. Xie, T. D. Pham, S. Shetty, F. A. Son, K. B. Idrees, Z. Chen, H. Xie, Y. Liu, R. Q. Snurr, B. Chen, B. Alameddine, Y. Cui and O. K. Farha, *J. Am. Chem. Soc.*, 2022, **144**, 3737–3745.
18 R.-B. Lin, Z. Zhang and B. Chen, *Acc. Chem. Res.*, 2021, **54**, 3362–3376.
19 P. Z. Moghadam, Y. G. Chung and R. Q. Snurr, *Nat. Energy*, 2024, **9**, 121–133.

20 X. Zhang, R.-B. Lin, Z. A. Alothman, O. Alduhaish, T. Yildirim, W. Zhou, J.-R. Li and B. Chen, *Inorg. Chem. Front.*, 2023, **10**, 454–459.
21 M. Chang, T. Yan, Y. Wei, J.-X. Wang, D. Liu and J.-F. Chen, *ACS Appl. Mater. Interfaces*, 2022, **14**, 25374–25384.
22 F. Gándara, H. Furukawa, S. Lee and O. M. Yaghi, *J. Am. Chem. Soc.*, 2014, **136**, 5271–5274.
23 L. Ali and E. Mahmoud, *J. Porous Mater.*, 2021, **28**, 213–230.
24 X. Li, X. Zhang, J. Zhang, J. Gu, S. Zhang, G. Li, J. Shao, Y. He, H. Yang, S. Zhang and H. Chen, *Carbon Capture Sci. Technol.*, 2023, **9**, 100146.
25 J. Behler, *Angew. Chem., Int. Ed.*, 2017, **56**, 12828–12840.
26 Y. Dong, Y. Liu, F. Mao and H. Wu, *Appl. Energy*, 2025, **393**, 126074.
27 M. Fernandez, P. G. Boyd, T. D. Daff, M. Z. Aghaji and T. K. Woo, *J. Phys. Chem. Lett.*, 2014, **5**, 3056–3060.
28 H. Wan, Y. Fang, M. Hu, S. Guo, Z. Sui, X. Huang, Z. Liu, Y. Zhao, H. Liang and Y. Wu, *Adv. Sci.*, 2025, **12**, 2405905.
29 R. Xin, C. Wang, Y. Zhang, R. Peng, R. Li, J. Wang, Y. Mao, X. Zhu, W. Zhu and M. Kim, *ACS Nano*, 2024, **18**, 19403–19422.
30 Z. Wang, T. Zhou and K. Sundmacher, *Chem. Eng. J.*, 2022, **444**, 136651.
31 S. Majumdar, S. M. Moosavi, K. M. Jablonka, D. Ongari and B. Smit, *ACS Appl. Mater. Interfaces*, 2021, **13**, 61004–61014.
32 K. M. Jablonka, D. Ongari, S. M. Moosavi and B. Smit, *Chem. Rev.*, 2020, **120**, 8066–8129.
33 W. E. Marcilio and D. M. Eler, in. *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, IEEE, 2020, pp. 340–347.
34 S. Meduri and J. Nandanavanam, *Energy AI*, 2023, **12**, 100230.
35 S. Namdeo, V. C. Srivastava and P. Mohanty, *J. Colloid Interface Sci.*, 2023, **647**, 174–187.
36 X. Wei, Z. Lu, Y. Ai, L. Shen, M. Wei and X. Wang, *Sep. Purif. Technol.*, 2024, **330**, 125291.
37 M. Zhong, K. Tran, Y. Min, C. Wang, Z. Wang, C.-T. Dinh, P. De Luna, Z. Yu, A. S. Rasouli, P. Brodersen, S. Sun, O. Voznyy, C.-S. Tan, M. Askerka, F. Che, M. Liu, A. Seifitokaldani, Y. Pang, S.-C. Lo, A. Ip, Z. Ulissi and E. H. Sargent, *Nature*, 2020, **581**, 178–183.
38 H. Liu, W. Ding, F. Zhou, G. Yang and Y. Du, *Sep. Sci. Technol.*, 2020, **55**, 1102–1114.
39 B. Bonelli, in. *Advances in natural gas: formation, processing, and applications. volume 7: natural gas products and uses*, Elsevier, 2024, pp. 401–432.
40 D.-A. Chisăliță, J. Boon and L. Lücking, *Sep. Purif. Technol.*, 2025, **353**, 128466.
41 D. Ursueguía, E. Díaz and S. Ordóñez, *Sci. Total Environ.*, 2021, **790**, 148211.
42 X. Wei, D. Peng, L. Shen, Y. Ai and Z. Lu, *J. Environ. Chem. Eng.*, 2023, **11**, 110086.
43 S. Lee, B. Kim, H. Cho, H. Lee, S. Y. Lee, E. S. Cho and J. Kim, *ACS Appl. Mater. Interfaces*, 2021, **13**, 23647–23654.
44 R. Anderson, A. Biong and D. A. Gómez-Gualdrón, *J. Chem. Theory Comput.*, 2020, **16**, 1271–1283.
45 R. Gurnani, Z. Yu, C. Kim, D. S. Sholl and R. Ramprasad, *Chem. Mater.*, 2021, **33**, 3543–3552.
46 Y. Zhang and Q. Yang, *IEEE Trans. Knowl. Data Eng.*, 2022, **34**, 5586–5609.
47 Y. Dai, Y. Xiong, H. Lin, Y. Li, Y. Feng, W. Luo and X. Zhong, *IEEE Sens. J.*, 2024, **24**, 11639–11650.
48 M. Yin, X. Zhang, F. Li, X. Yan, X. Zhou, Q. Ran, K. Jiang, T. Borch and L. Fang, *Environ. Sci. Technol.*, 2023, **58**, 1771–1782.
49 Y. J. Colón, D. A. Gómez-Gualdrón and R. Q. Snurr, *Cryst. Growth Des.*, 2017, **17**, 5801–5810.
50 Y. G. Chung, D. A. Gómez-Gualdrón, P. Li, K. T. Leperi, P. Deria, H. Zhang, N. A. Vermeulen, J. F. Stoddart, F. You, J. T. Hupp, O. K. Farha and R. Q. Snurr, *Sci. Adv.*, 2016, **2**, e1600909.
51 J. Burner, J. Luo, A. White, A. Mirmiran, O. Kwon, P. G. Boyd, S. Maley, M. Gibaldi, S. Simrod, V. Ogden and T. K. Woo, *Chem. Mater.*, 2023, **35**, 900–916.
52 A. White, M. Gibaldi, J. Burner, R. A. Mayo and T. Woo, *ChemRxiv*, 2024, DOI: 10.26434/chemrxiv-2024-ftsv3.
53 T. F. Willems, C. H. Rycroft, M. Kazi, J. C. Meza and M. Haranczyk, *Microporous Mesoporous Mater.*, 2012, **149**, 134–141.
54 Y. Liu, Y. Dong and H. Wu, *J. Mater. Chem. A*, 2025, **13**, 2403–2440.
55 J. P. Janet and H. J. Kulik, *J. Phys. Chem. A*, 2017, **121**, 8939–8954.
56 S. M. Moosavi, A. Nandy, K. M. Jablonka, D. Ongari, J. P. Janet, P. G. Boyd, Y. Lee, B. Smit and H. J. Kulik, *Nat. Commun.*, 2020, **11**, 4068.
57 J. P. Janet, F. Liu, A. Nandy, C. Duan, T. Yang, S. Lin and H. J. Kulik, *Inorg. Chem.*, 2019, **58**, 10592–10606.
58 D. Dubbeldam, S. Calero, D. E. Ellis and R. Q. Snurr, *Mol. Simul.*, 2016, **42**, 81–101.
59 J. M. Stubbs, J. J. Potoff and J. I. Siepmann, *J. Phys. Chem. B*, 2004, **108**, 17596–17605.
60 A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.
61 S. Vandenhende, S. Georgoulis, W. Van Gansbeke, M. Proesmans, D. Dai and L. Van Gool, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2021, 1.
62 R. Cipolla, Y. Gal and A. Kendall, in. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, 2018, pp. 7482–7491.
63 S. Lundberg and S.-I. Lee, *Adv. Neural Inf. Process. Syst.*, 2017, **30**, 4765–4774.
64 T. A. Makal, J.-R. Li, W. Lu and H.-C. Zhou, *Chem. Soc. Rev.*, 2012, **41**, 7761–7779.
65 T. Wang, Q. Zhang, Z. Zhao, L. Zhang and B. Xing, *Chem. Eng. J.*, 2024, **486**, 150150.
66 M. Witman, S. Ling, V. Stavila, P. Wijeratne, H. Furukawa and M. D. Allendorf, *Mol. Syst. Des. Eng.*, 2020, **5**, 1491–1503.
67 I. Melián-Cabrera and V. Zarubina, *Microporous Mesoporous Mater.*, 2023, **354**, 112549.
68 X. Wu, S. Xiang, J. Su and W. Cai, *J. Phys. Chem. C*, 2019, **123**, 8550–8559.
69 Y. Han, J. Zhao, X. Guo and T. Jiao, *Langmuir*, 2023, **39**, 7648–7659.

70 H. Long, H. Lin, M. Yan, Y. Bai, X. Tong, X. Kong and S. Li, *Fuel*, 2021, **292**, 120268.
71 B. Li, H.-M. Wen, H. Wang, H. Wu, M. Tyagi, T. Yildirim, W. Zhou and B. Chen, *J. Am. Chem. Soc.*, 2014, **136**, 6207–6210.
72 B. Li, H.-M. Wen, H. Wang, H. Wu, T. Yildirim, W. Zhou and B. Chen, *Energy Environ. Sci.*, 2015, **8**, 2504–2511.
73 Y. Peng, V. Krungleviciute, I. Eryazici, J. T. Hupp, O. K. Farha and T. Yildirim, *J. Am. Chem. Soc.*, 2013, **135**, 11887–11894.
74 H. Furukawa, N. Ko, Y. B. Go, N. Aratani, S. B. Choi, E. Choi, A. Ö. Yazaydin, R. Q. Snurr, M. O'Keeffe and J. Kim, *Science*, 2010, **329**, 424–428.
75 T. N. Tu, H. T. D. Nguyen and N. T. Tran, *Inorg. Chem. Front.*, 2019, **6**, 2441–2447.