


 Cite this: *Mol. Syst. Des. Eng.*, 2025, **10**, 1005

# Computer-aided metal–organic framework screening and design approaches toward efficient carbon capture processes

 Zihao Wang <sup>a</sup> and Teng Zhou <sup>\*bc</sup>

Carbon capture is a priority strategy for reducing CO<sub>2</sub> emissions and mitigating climate change. Adsorption-based technologies offer significant potential to reduce imposed parasitic energy, and metal–organic frameworks (MOFs) are considered a promising class of adsorbents for this purpose. In this review, targeting carbon capture using MOFs, we explore materials screening approaches using material-level properties (e.g., CO<sub>2</sub> working capacity and CO<sub>2</sub>/N<sub>2</sub> selectivity) and process-level performance indicators (e.g., CO<sub>2</sub> purity and energy consumption), with an emphasis on the incorporation of process-level considerations into screening workflows. We also highlight recent advancements of data-driven property and process models in accelerating large-scale materials screening. Next, we review diverse materials design approaches, shifting from open-loop exhaustive search to closed-loop targeted discovery. Finally, we discuss the challenges associated with experimental databases, active materials discovery, and simultaneous material and process design, with perspectives proposed to accelerate the materials discovery for industrial carbon capture applications.

 Received 29th April 2025,  
 Accepted 2nd September 2025

DOI: 10.1039/d5me00075k

[rsc.li/molecular-engineering](https://rsc.li/molecular-engineering)

## Design, System, Application

Carbon capture is a critical strategy for mitigating global climate change and achieving a net-zero future. Significant research efforts have been dedicated to developing advanced adsorbents, such as metal–organic frameworks (MOFs), with improved properties for CO<sub>2</sub> capture. Traditionally, this process is typically dominated by trial-and-error methods, which are resource-intensive and heavily depend on expert knowledge. In contrast, computational approaches are increasingly being employed to evaluate adsorbents, enabling targeted experimental validation of promising candidates. Targeting efficient carbon capture using MOFs, this review explores computer-aided materials screening and design approaches. Key computational techniques including molecular simulations, structure–property relationships, and process simulation and optimization are comprehensively introduced. Multi-scale screening approaches spanning from material to process level are discussed, highlighting the importance of process-level considerations in identifying high-performance MOFs for carbon capture. Additionally, recent materials design strategies are introduced for tailoring MOFs with desired properties, shifting from open-loop exhaustive search to closed-loop targeted discovery. Emerging machine learning and data-driven techniques for accelerated materials discovery are also explored. Overall, this review presents a comprehensive overview of innovative and interdisciplinary approaches for the efficient discovery of high-performance MOFs for carbon capture, which are also broadly applicable to a wide range of adsorbent materials and separation applications.

## 1. Introduction

Reducing carbon dioxide (CO<sub>2</sub>) emissions from industry, transportation, buildings, and other sources is a critical priority for mitigating global climate change and achieving a net-zero future. Carbon capture, deployed at an industrial scale, is a particularly attractive and promising option.<sup>1</sup> Various carbon capture technologies such as post-

combustion capture can be retrofitted onto existing power plants, offering a practical solution for reducing CO<sub>2</sub> emissions in short and medium terms.<sup>2</sup> However, the main challenge in the implementation of carbon capture technologies is the significant additional energy costs associated with the capture process.<sup>3</sup> Consequently, developing energy-efficient capture processes that minimize the energy consumption imposed by carbon capture has become an important area of research.

Traditional approaches for carbon capture from power plant streams primarily rely on amine-based absorption processes.<sup>2,4</sup> In amine scrubbing, CO<sub>2</sub> is absorbed by an aqueous amine solution, followed by stripping of the captured CO<sub>2</sub> out of the liquid phase through heating.<sup>5</sup> Although this process is highly efficient due to the high CO<sub>2</sub>

<sup>a</sup> School of Chemistry and Chemical Engineering, Chongqing University, Chongqing 401331, China

<sup>b</sup> Sustainable Energy and Environment Thrust, The Hong Kong University of Science and Technology (Guangzhou), Nansha, Guangzhou 511400, China.

E-mail: [tengzhou@hkust-gz.edu.cn](mailto:tengzhou@hkust-gz.edu.cn)

<sup>c</sup> Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China



selectivity of amine solutions, a significant energy penalty associated with heating the aqueous amine is imposed, leading to high operating costs.<sup>6,7</sup> In this context, alternative technologies, such as adsorption-based processes using solid adsorbents, have been considered for their potential to reduce energy consumption in carbon capture.

Adsorption processes have been widely used for gas separations, such as air separation and hydrogen purification.<sup>8</sup> For carbon capture, adsorption has emerged as a promising alternative to traditional amine-based processes due to its potential to lower regeneration costs. Typical adsorption processes include pressure swing adsorption (PSA), vacuum swing adsorption (VSA), and temperature swing adsorption (TSA).<sup>9</sup> The effectiveness and efficiency of these processes are highly dependent on the adsorbent material, and the development of improved adsorbents remains a central challenge for realizing the industrial-scale deployment of adsorption-based carbon capture technologies.

In recent years, various types of solid adsorbents including porous carbons, zeolites, covalent organic frameworks, and metal-organic frameworks (MOFs), have been actively synthesized and evaluated for adsorption-based carbon capture.<sup>10-12</sup> Among these available materials, MOFs are highly attractive due to their vast chemical design space. In principle, by combining different metal nodes and organic linkers, numerous structurally diverse MOFs with distinct physical and chemical properties can be engineered.<sup>13</sup> This tunability and versatility make MOFs ideal adsorbents for a wide range of applications. They have also been demonstrated as promising candidates for carbon capture applications due to their high CO<sub>2</sub> working capacity and selectivity over other components. For example, a recent innovative zinc-based MOF, CALF-20,<sup>14</sup> has demonstrated

remarkable stability in humid conditions and high selectivity to capture CO<sub>2</sub> from flue gas. It can also meet all criteria required for post-combustion CO<sub>2</sub> capture, offering a cost-effective and reliable option for industrial-scale applications.

Huge research efforts have been dedicated to developing better adsorbents with improved properties such as higher working capacity and selectivity.<sup>1</sup> This is typically dominated by traditional trial-and-error strategies, heavily dependent on the domain knowledge and expertise of chemists and materials scientists.<sup>15</sup> Thus, this process is often resource-intensive and time-consuming. Alternatively, computational methods have been increasingly adopted to evaluate the adsorbent's performance, followed by experimental validation focusing only on the most promising candidates.<sup>16</sup> Furthermore, computational methods allow researchers to explore the potential performance of adsorbent materials that have not yet been synthesized. This enables the discovery of new materials with improved performance, as well as the identification of existing materials suitable for novel applications. Therefore, computer-aided screening and design have become essential tools for identifying high-performance adsorbents from large materials databases and accelerating the discovery of new adsorbents for carbon capture.

In this review, targeting carbon capture using MOFs, we explore the computer-aided materials screening approaches from a multi-scale perspective, spanning from the material to the process level. We also introduce computer-aided materials design strategies for tailoring MOFs with desired properties, shifting from open-loop exhaustive search to closed-loop targeted discovery. In light of significant advancements in machine learning and related fields, we extend the scope of this review to include their key applications in accelerating materials screening and design.



**Zihao Wang**

*Zihao Wang is a Research Associate at Chongqing University. He earned his Ph.D. in Process Systems Engineering in 2025 through the International Max Planck Research School for Advanced Methods in Process and Systems Engineering (IMPRS-ProEng), a joint program of the Max Planck Institute for Dynamics of Complex Technical Systems and the Otto von Guericke University Magdeburg. He holds both*

*bachelor's and master's degrees in Chemical Engineering and Technology from Chongqing University, completed in 2017 and 2020, respectively. His research interests include data-driven multiscale modeling and optimization, efficient molecular and materials discovery, and design and optimization of complex separation and energy systems.*



**Teng Zhou**

*Teng Zhou is an Assistant Professor and Head of the Max Planck Partner Group at the Hong Kong University of Science and Technology (Guangzhou). He received his Ph.D. with highest distinction in 2016 from the Max Planck Institute for Dynamics of Complex Technical Systems. Dr. Zhou's research focuses on advanced materials and process design for CO<sub>2</sub> capture, integrated capture and conversion, and energy systems*

*engineering. He has published 70 journal articles and serves on editorial boards of several leading journals. He has been recognized with prestigious honors, including the Carbon Capture Future Leader Award and the AIChE CAST Directors Award.*



Although the primary focus is on MOFs and carbon capture applications, the methodologies and challenges discussed are broadly applicable to a wide range of adsorbent materials (e.g., zeolites and covalent organic frameworks) and separation applications (e.g., hydrogen purification and hydrocarbon separation).

## 2. Computer-aided screening of metal–organic frameworks

To date, tens of thousands of MOF structures have been synthesized and evaluated for various applications, and open-source structure databases have been established, serving as essential resources for advancing computational materials science and related fields. One of the largest and most widely used databases is the hypothetical MOF (hMOF) database, established by Wilmer *et al.*<sup>17</sup> in 2012. This database contains 137 953 structures generated by combining available metal nodes and organic linkers from experimentally synthesized MOFs. However, its structural diversity is relatively limited, as only a few metal nodes and framework topologies were considered. A significant advancement in computational materials science came with the establishment of the Computation-Ready, Experimental (CoRE) MOF database in 2014.<sup>18</sup> This database initially compiled over 4700 experimentally reported MOFs, and these structures were modified to be directly used in atomistic simulations, offering a variety of experimentally synthesizable MOF structures for preliminary computational investigations. In 2019, it was expanded to include approximately 14 000 structures, leading to the release of the CoRE MOF 2019 database.<sup>19</sup> The database was recently updated to include over 40 000 structures reported by early 2024, along with a comprehensive set of calculated material properties.<sup>20</sup> Most of these major experimental and hypothetical structure databases have been integrated into the online platform MOFX-DB,<sup>21</sup> which provides simulated adsorption data for gases (e.g., H<sub>2</sub>, CH<sub>4</sub>, CO<sub>2</sub>, and N<sub>2</sub>), textural properties (e.g., pore sizes and surface areas), and structure files, enabling a user-oriented search for MOF structures with desired structural and adsorption characteristics. In addition to hMOF and CoRE MOF databases, a handful of databases of varying sizes and resources, such as Cambridge Structural Database (CSD) and QMOF, have been developed.<sup>22–28</sup> These open-source resources have significantly advanced computational materials science, offering extensive options for diverse MOF-supported applications including carbon capture.

To identify promising candidates from MOF databases, systematic and efficient methods for large-scale screening are highly desirable. Computational tools such as Zeo++<sup>29</sup> and Poreblazer<sup>30</sup> have been developed to obtain key structural characteristics (e.g., surface area and porosity). When combined with experimental expertise, these structural characteristics can be used to guide the selection of materials. However, they do not capture the interaction

between the material and its environment, which is crucial for separation applications such as carbon capture, where competitive adsorption of different components (e.g., CO<sub>2</sub>, N<sub>2</sub>, and H<sub>2</sub>O) on the material must be considered. Therefore, relying solely on these structural characteristics for materials selection may not be effective. In such cases, suitable metrics that can quantify the performance of materials are necessary to identify optimal candidates from MOF databases.

In general, metrics for materials screening can be classified into two types: material- and process-level metrics. Material-level metrics quantify the ability of materials to adsorb and separate different components under specific conditions, while process-level metrics evaluate the ability of materials to separate different components in practical adsorption processes. This section focuses on materials screening methods for MOFs, transitioning from property-based screening using material-level metrics (e.g., CO<sub>2</sub> working capacity and CO<sub>2</sub>/N<sub>2</sub> selectivity) to performance-based screening that incorporates process-level considerations (e.g., CO<sub>2</sub> purity and energy consumption) (see Fig. 1).

### 2.1. Material level: property-based screening

Over the years, a wide range of material-level metrics has been developed, primarily based on adsorption equilibrium data. The most common and straightforward metrics include selectivity, which indicates the relative affinities of two components on an adsorbent, and working capacity, which quantifies the difference in equilibrium adsorption capacities between adsorption and regeneration conditions.<sup>31</sup> More advanced material-level metrics include adsorption selectivity,<sup>32</sup> separation factor,<sup>33</sup> regenerability,<sup>32</sup> and adsorbent performance indicator,<sup>34</sup> among others. For example, the adsorbent performance score (APS) was introduced to combine both selectivity and working capacity, as materials with high working capacity usually exhibit low selectivity, and *vice versa*.<sup>35</sup> A comprehensive list of these material-level metrics and their definition are available in the literature.<sup>2,6</sup> These metrics typically rely on single- and multi-component adsorption isotherms, with some also incorporating the enthalpy of adsorption to account for the ease of adsorbent regeneration. With significant advancements in computational power and algorithms, adsorption isotherms and enthalpies of adsorption can be efficiently calculated, enabling large-scale screening of materials databases to identify promising candidates for carbon capture based on material-level metrics.

**2.1.1. Molecular modeling.** Since experimental adsorption data is available for only a limited number of materials, molecular modeling methods such as quantum mechanical calculations and molecular simulations provide a powerful alternative for estimating gas adsorption behaviors across diverse materials.<sup>36</sup> Using these computational methods, various adsorption-related properties can be determined to gain key insights into material performance in carbon



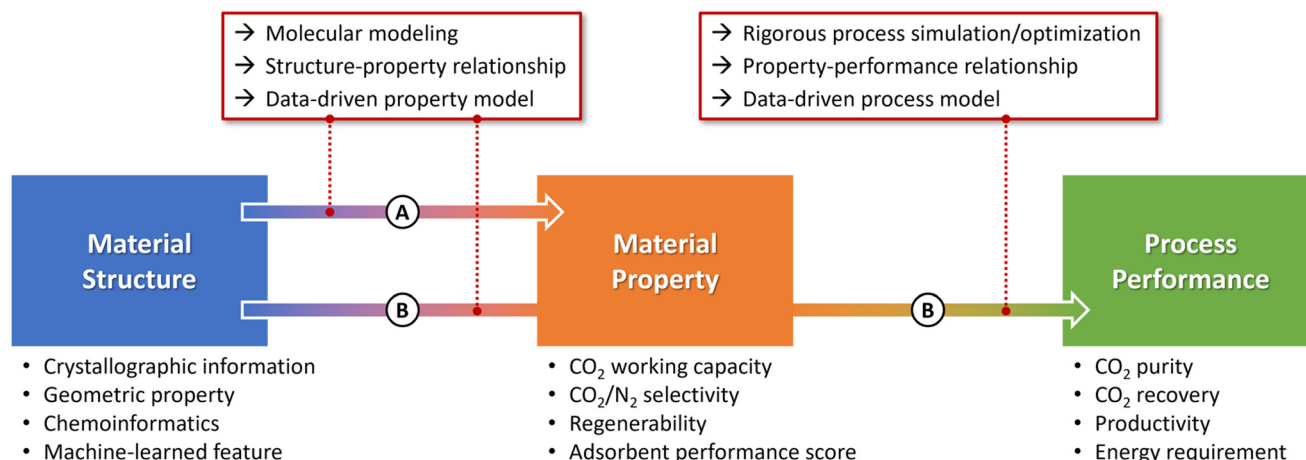


Fig. 1 Structure–property–performance hierarchy and general routines for materials screening in carbon capture: (A) material-level property-based screening and (B) process-level performance-based screening.

capture. These properties include thermodynamic properties (e.g., CO<sub>2</sub> adsorption isotherms and enthalpies of CO<sub>2</sub> adsorption) and kinetic properties (e.g., diffusion coefficients of CO<sub>2</sub>). Compared to quantum mechanical calculations such as density functional theory (DFT), molecular simulations are less computationally demanding and therefore are relatively affordable for screening hundreds to thousands of materials. For example, grand canonical Monte Carlo (GCMC) simulations are widely used to estimate equilibrium adsorption capacities and enthalpies of adsorption, enabling large-scale materials screening based on material-level metrics.<sup>37</sup>

To identify suitable MOF structures for precombustion CO<sub>2</sub> capture (i.e., separating CO<sub>2</sub> from syngas), Avci *et al.*<sup>38</sup> evaluated 3857 diverse MOF structures from the CSD database. They performed GCMC simulations for CO<sub>2</sub>/H<sub>2</sub> mixtures at 0.1, 1, and 10 bar, and key material-level metrics including CO<sub>2</sub>/H<sub>2</sub> selectivity, CO<sub>2</sub> working capacity, APS, and regenerability were calculated for different adsorbent regeneration methods. Under the constraint that regenerability should exceed 85%, they identified 20 promising MOFs with the highest APS values. One of the MOFs identified was previously reported with CO<sub>2</sub> and H<sub>2</sub> adsorption capacities. Further validation under the identical conditions confirmed a good agreement between the GCMC simulations and the reported experimental data. This demonstrated that molecular simulations accurately estimated MOF properties, and materials screening based on the APS successfully identified high-performing candidates from the large set of materials. In a more recent study by the same group,<sup>39</sup> they expanded the candidate pool to 10 221 MOFs using the updated CSD database. This broader materials screening enabled the identification of MOFs with both higher selectivity and CO<sub>2</sub> working capacity.

For post-combustion CO<sub>2</sub> capture (i.e., separating CO<sub>2</sub> from flue gas) and natural gas upgrading (i.e., removing CO<sub>2</sub> from natural gas), Qiao *et al.*<sup>40</sup> used molecular simulations to evaluate 4764 MOFs from the CoRE MOF 2014 database. GCMC

simulations were performed for CO<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub>/CH<sub>4</sub> mixtures. Material performance was assessed using four material-level metrics, including CO<sub>2</sub> selectivity, CO<sub>2</sub> adsorption capacity, CO<sub>2</sub> working capacity, and regenerability. Under specified constraints on CO<sub>2</sub> working capacity, CO<sub>2</sub> selectivity, and regenerability, 15 top-performing MOFs were identified. Further breakthrough simulations confirmed the high performance of these adsorbents in separating CO<sub>2</sub> from flue gas or natural gas, demonstrating the effectiveness of molecular simulations in identifying promising materials for diverse carbon capture scenarios.

For 4188 MOFs selected from the CoRE MOF 2019 database, Xu *et al.*<sup>41</sup> evaluated their performance in post-combustion CO<sub>2</sub> capture *via* GCMC simulations. MOFs with high CO<sub>2</sub> working capacity, low N<sub>2</sub> working capacity, high CO<sub>2</sub> selectivity, and high regenerability were retained, leading to 90 MOFs being identified as superior adsorbents for the CO<sub>2</sub>/N<sub>2</sub> separation. Further analysis of adsorption isosteric enthalpy revealed that these MOFs exhibit stronger interactions with CO<sub>2</sub> than N<sub>2</sub> molecules, confirming their effectiveness in selective CO<sub>2</sub> adsorption. Furthermore, when evaluated across varying CO<sub>2</sub> concentrations in flue gas, some of these MOFs exhibited adaptability, which makes them suitable candidates for industrial carbon capture applications.

When applied to carbon capture for water-containing gas mixtures (e.g., wet flue gas), many MOFs exhibit reduced performance because water competes with CO<sub>2</sub> for adsorption sites.<sup>42</sup> To address this challenge, systematic screening approaches that consider the effect of water are essential for identifying materials suitable for carbon capture under humid conditions. Kancharlapalli and Snurr<sup>43</sup> proposed a multi-scale modeling strategy to screen MOFs for selective CO<sub>2</sub> adsorption from wet flue gas. For 3703 MOFs selected from the CoRE MOF 2019 database, they optimized single-molecule adsorption configurations and calculated MOF-molecule interaction energies for CO<sub>2</sub>, N<sub>2</sub>, and H<sub>2</sub>O. This initial screening identified 458 MOFs with a stronger affinity for CO<sub>2</sub> than for both N<sub>2</sub> and H<sub>2</sub>O. DFT calculations



were performed to accurately determine adsorption energies between molecules and MOFs, and 48 MOFs were confirmed to have stronger interactions with CO<sub>2</sub> than with H<sub>2</sub>O. Subsequently, GCMC simulations of a binary CO<sub>2</sub>/N<sub>2</sub> mixture narrowed the list to 10 MOFs that exhibit high CO<sub>2</sub>/N<sub>2</sub> selectivity and CO<sub>2</sub> adsorption capacity. Finally, in ternary mixture simulations considering the presence of water, most of these MOFs are still highly selective for CO<sub>2</sub> adsorption. This demonstrated the effectiveness of the multi-scale screening approach in identifying promising MOFs for selective CO<sub>2</sub> capture from wet flue gas.

For natural gas upgrading in the presence of water, Rogacka *et al.*<sup>44</sup> initially selected 764 MOFs from the CoRE MOF 2014-DDEC database<sup>24</sup> using pore limiting diameter and Henry constant of water adsorption as screening criteria. Subsequent GCMC simulations for the ternary CO<sub>2</sub>/CH<sub>4</sub>/H<sub>2</sub>O mixture under different humidity levels identified 13 MOFs with high adsorption capacity and selectivity over H<sub>2</sub>O.

Li *et al.*<sup>45</sup> conducted large-scale computational screening to identify suitable MOFs that can effectively adsorb CO<sub>2</sub> at 80% relative humidity for post-combustion CO<sub>2</sub> capture. Starting with 5109 MOFs from the CoRE MOF 2014 database, they removed structures with zero accessible surface area and high H<sub>2</sub>O adsorption energy, reducing the pool to 2054 MOFs. These remaining structures were then ranked by CO<sub>2</sub>/H<sub>2</sub>O Henry's selectivity, and the top 15 MOFs were selected for further binary (CO<sub>2</sub>/H<sub>2</sub>O) and ternary (CO<sub>2</sub>/H<sub>2</sub>O/N<sub>2</sub>) GCMC simulations. The CO<sub>2</sub>/H<sub>2</sub>O selectivity obtained from both binary and ternary simulations showed a strong agreement. These top-performing MOFs exhibited high CO<sub>2</sub>/H<sub>2</sub>O and CO<sub>2</sub>/N<sub>2</sub> selectivity, confirming their potential for highly selective post-combustion CO<sub>2</sub> capture under humid conditions.

**2.1.2. Structure–property relationship.** Computationally expensive molecular simulations are usually required to evaluate each candidate material of interest, posing a major challenge for large-scale screening. To address this, property models that correlate material structures with their properties offer efficient alternatives for adsorption property estimation. However, unlike small molecules, materials lack well-developed property models due to their structural complexity. In this context, various mathematical modeling techniques, including machine learning approaches, have been used to establish structure–property relationship models for materials.<sup>46,47</sup> These models, developed on relatively small datasets, can rapidly predict targeted properties for a large number of candidate materials. In this way, only the most promising materials were subsequently evaluated using molecular simulations, significantly accelerating large-scale, property-based materials screening.

An important consideration in establishing structure–property relationships is the representation of materials, which encodes material structures into machine-readable formats.<sup>48</sup> The most fundamental representations are textural properties, such as pore volume and surface area, which directly describe the pore geometry of material structures.<sup>49</sup>

Beyond these geometry-related properties, chemical information can be obtained from material structures as descriptors, such as the metal node and the chemical composition of organic linkers.<sup>50,51</sup> Additionally, energy-based descriptors can be derived from molecular modeling to account for the interaction behavior between adsorbates and MOFs.<sup>52,53</sup> In a more advanced manner, various algorithms were proposed to extract structural features, which are usually not readily interpretable by humans. Despite this, these approaches can encode comprehensive structural and chemical information into high-dimensional representations, leading to accurate predictions of material properties. For example, graph neural networks have been employed to learn hierarchical features from both building blocks and entire reticular crystal structures.<sup>54–57</sup>

Using five fundamental geometric features (such as pore diameter, void fraction, and surface area), Fernandez and Barnard<sup>58</sup> developed a set of models (such as multilinear regression and artificial neural network) to classify MOFs as high- and low-performing based on their CO<sub>2</sub> and N<sub>2</sub> adsorption capacities. Among these models, the random forest (RF) model showed the highest accuracy over 94% for both capacities. When applied to a large and chemically diverse set of approximately 65 000 MOFs, the model successfully identified over 70% of high-performing MOFs for CO<sub>2</sub> capture and 60% for N<sub>2</sub> capture. By pre-screening MOFs using the RF model, the pool of candidate materials can be considerably reduced, allowing computational resources to be focused on the most promising materials.

Combining six geometric features and three chemical features derived from the atomic property weighted radial distribution function,<sup>59</sup> Dureckova *et al.*<sup>49</sup> trained gradient boosted trees regression (GBTR) models to predict CO<sub>2</sub> working capacity and CO<sub>2</sub>/H<sub>2</sub> selectivity of MOFs for precombustion CO<sub>2</sub> capture. These models presented high accuracy for both CO<sub>2</sub> working capacity and CO<sub>2</sub>/H<sub>2</sub> selectivity, with determination coefficients of 0.944 and 0.872, respectively. When applied to a dataset of 35 840 MOFs, these models correctly identified 77.1% of the top 1000 MOFs for CO<sub>2</sub> working capacity and 86.0% for CO<sub>2</sub>/H<sub>2</sub> selectivity. These results indicate that these accurate models can significantly accelerate the screening of top-performing MOFs for precombustion CO<sub>2</sub> capture.

In addition to these, a variety of models for structure–property relationships have been developed to accurately and efficiently predict key material properties (*e.g.*, CO<sub>2</sub> working capacity and selectivity) relevant to carbon capture applications.<sup>60–62</sup> These models provide a wide range of predictive tools for large-scale property-based materials screening, enabling rapid identification of promising MOF structures for specific carbon capture applications.

## 2.2. Process level: performance-based screening

Materials screening based on material-level metrics typically overlooks the dynamics of adsorption processes, and



therefore, the screening outcomes may not accurately reflect material performance under practical operational conditions.<sup>63</sup> To bridge this gap, process modeling needs to be directly integrated into the materials screening workflow to evaluate the material's performance in dynamic adsorption processes such as PSA, VSA, and TSA. Relying on equilibrium data (*e.g.*, adsorption isotherms), dynamic properties (*e.g.*, diffusivity), and other characteristics (*e.g.*, thermal properties), process simulation and optimization can be conducted for adsorbent materials to obtain their process-level metrics such as CO<sub>2</sub> purity and recovery, adsorbent usage, and energy consumption. These metrics not only reflect the realistic process performance but also directly relate to the techno-economic analysis of adsorption processes in carbon capture. Therefore, performance-based approaches can shift the focus of materials screening from equilibrium properties at the material level to separation performance at the process

level, providing a more practical way to identify high-performance MOFs for carbon capture applications.

**2.2.1. Rigorous process model.** Adsorption processes are typically described by a set of mass, energy, and momentum balance equations, coupled with mass transfer and adsorption isotherm equations. Integration of these physics-based equations leads to a system of partial differential equations (PDEs), which needs to be solved to reveal the dynamic profiles of adsorption processes over both space and time. Relying on this PDE system, process optimization can be performed to determine the optimal performance of adsorbent materials across a wide range of operating conditions while simultaneously satisfying desired specifications such as the purity requirement, offering process-level evaluation and screening of materials.

Several well-known materials have been synthesized and validated as promising adsorbents for carbon capture, and their process-level metrics can be directly evaluated using

	WC	$\alpha$	S	API <sub>1</sub>	API <sub>2</sub>	AFM <sub>1</sub>	AFM <sub>2</sub>	SF	GEM	Mod. Skars.	5-Step Cycle	FVSA
1	Ni-MOF-74	SIFSIX-3-Ni	SIFSIX-3-Ni	Mg-MOF-74	Mg-MOF-74	SIFSIX-3-Ni	SIFSIX-3-Ni	SIFSIX-3-Ni	UTSA-16	Zeolite 13X	UTSA-16	Cu-TDPAT
2	Mg-MOF-74	Zeolite 13X	Zeolite 13X	Zeolite 13X	Ni-MOF-74	Zeolite 13X	Zeolite 13X	UTSA-16	Cu-TDPAT	Cu-TDPAT	Zeolite 13X	UTSA-16
3	Co-MOF-74	Mg-MOF-74	UTSA-16	SIFSIX-3-Ni	UTSA-16	Mg-MOF-74	UTSA-16	Zeolite 13X	Zeolite 13X	UTSA-16	Cu-TDPAT	Zeolite 13X
4	UiO-66(OH) <sub>2</sub>	UTSA-16	Cu-TDPAT	Ni-MOF-74	Zeolite 13X	Ni-MOF-74	Cu-TDPAT	Ni-MOF-74	SIFSIX-3-Ni	Ni-MOF-74	Ti-MIL-91	Zn-MOF-74
5	UTSA-16	Cu-TDPAT	Ni-MOF-74	UTSA-16	SIFSIX-2-Cu-i	UTSA-16	Ni-MOF-74	Cu-TDPAT	Ti-MIL-91	Mg-MOF-74	Zn-MOF-74	Ti-MIL-91
6	SIFSIX-2-Cu-i	Ni-MOF-74	Mg-MOF-74	Cu-TDPAT	Co-MOF-74	Cu-TDPAT	Mg-MOF-74	Ti-MIL-91	Zn-MOF-74	SIFSIX-2-Cu-i	SIFSIX-3-Ni	SIFSIX-3-Ni
7	Zn-MOF-74	Ti-MIL-91	Ti-MIL-91	SIFSIX-2-Cu-i	UiO-66(OH) <sub>2</sub>	SIFSIX-2-Cu-i	Ti-MIL-91	Mg-MOF-74	SIFSIX-2-Cu-i	SIFSIX-3-Ni	Mg-MOF-74	
8	Zeolite 13X	SIFSIX-2-Cu-i	SIFSIX-2-Cu-i	Zn-MOF-74	Zn-MOF-74	Zn-MOF-74	SIFSIX-2-Cu-i	SIFSIX-2-Cu-i	Ni-MOF-74	Zn-MOF-74		
9	Cu-TDPAT	Zn-MOF-74	Zn-MOF-74	UiO-66(OH) <sub>2</sub>	Cu-TDPAT	Ti-MIL-91	Zn-MOF-74	Zn-MOF-74	Sc2BDC3	Ti-MIL-91		
10	Ti-MIL-91	Sc2BDC3	Sc2BDC3	Ti-MIL-91	Ti-MIL-91	UiO-66(OH) <sub>2</sub>	Sc2BDC3	Sc2BDC3	Mg-MOF-74			
11	Cu-BTTRi	UiO-66(OH) <sub>2</sub>	UiO-66(OH) <sub>2</sub>	Co-MOF-74	SIFSIX-3-Ni	Co-MOF-74	UiO-66(OH) <sub>2</sub>	UiO-66(OH) <sub>2</sub>	Cu-BTTRi			
12	NTU-105	Co-MOF-74	Co-MOF-74	Cu-BTTRi	Cu-BTTRi	Cu-BTTRi	Co-MOF-74	Co-MOF-74	UiO-66(OH) <sub>2</sub>			
13	SIFSIX-3-Ni	Cu-BTTRi	Cu-BTTRi	NTU-105	NTU-105	NTU-105	Cu-BTTRi	Cu-BTTRi	ZIF-8			
14	MOF-177	NTU-105	NTU-105	Sc2BDC3	MOF-177	Sc2BDC3	NTU-105	NTU-105	Co-MOF-74			
15	ZIF-8	ZIF-8	ZIF-8	MOF-177	Sc2BDC3	MOF-177	ZIF-8	ZIF-8	NTU-105			
16	Sc2BDC3	MOF-177	MOF-177	ZIF-8	ZIF-8	ZIF-8	MOF-177	MOF-177	MOF-177			

**WC:** CO<sub>2</sub> working capacity  
 **$\alpha$ :** CO<sub>2</sub> selectivity  
**S:** sorbent selection parameter  
**API:** adsorbent performance indicator  
**AFM:** adsorbent figure of merit  
**SF:** separation factor  
**GEM:** general evaluation metric  
**Modified Skarstrom cycle**  
**Five-step cycle**  
**Fractionated vacuum swing adsorption cycle**

Fig. 2 Ranking of MOFs for post-combustion CO<sub>2</sub> capture using material-level properties and process-level performances. Used with permission of the Royal Society of Chemistry, from Yancy-Caballero et al.,<sup>64</sup> permission conveyed through Copyright Clearance Center, Inc.



detailed process models. For example, Yancy-Caballero *et al.*<sup>64</sup> performed simulation and optimization of PSA processes for 15 MOFs reported in the literature in post-combustion CO<sub>2</sub> capture. Using three different PSA cycle configurations, these MOFs were ranked based on their process performance such as productivity and energy requirements. This process-level ranking was then compared with rankings based on various material-level metrics, revealing that material-level metrics can sometimes lead to misleading evaluation of practical performance (Fig. 2). For instance, SIFSIX-3-Ni was identified as the best adsorbent by many material-level metrics, whereas it was ranked at the bottom in all three PSA processes. This highlights the importance of process-level considerations in screening high-performance materials for efficient carbon capture processes.

Pai *et al.* evaluated five MOFs for post-combustion CO<sub>2</sub> capture using a four-step VSA cycle with light product pressurization.<sup>65</sup> Through rigorous multi-objective process optimization, they identified two MOFs that outperformed the benchmark adsorbent, zeolite 13X, in terms of CO<sub>2</sub> purity and recovery. Further optimization of energy consumption and productivity demonstrated that both MOFs can reduce the energy required for CO<sub>2</sub> capture while achieving the required purity and recovery.

In the context of large materials databases, multi-scale screening approaches are typically employed. Initially, molecular modeling techniques are used to narrow down the pool of candidates, and the promising materials identified are then evaluated with detailed process simulation and optimization. For example, Park *et al.*<sup>66</sup> integrated molecular simulation, shortcut process modeling, and rigorous process optimization to identify promising MOFs for post-combustion CO<sub>2</sub> capture. Initially, they screened out 35 candidates from a pool of 143 MOFs based on material-level metrics including CO<sub>2</sub> working capacity, CO<sub>2</sub>/N<sub>2</sub> selectivity, and regenerability. This list was further reduced to 20 MOFs through PSA process modeling using shortcut models. Multi-objective optimization with rigorous process models was finally performed to rank these MOFs based on process-level metrics including CO<sub>2</sub> purity, recovery, productivity, and energy consumption. Among the material-level metrics considered, only the CO<sub>2</sub> working capacity presented a good correlation with process-level metrics in MOF rankings. Moreover, in comparison with rigorous process models, the shortcut models are too simplistic to provide sufficient process-level information, highlighting the importance of rigorous process models in reliable materials screening.

Burns *et al.*<sup>67</sup> integrated molecular simulations and rigorous process optimization to evaluate the performance of MOFs in an industrial VSA system for post-combustion CO<sub>2</sub> capture. After excluding structures with rare or toxic metals, 1584 MOFs from the CoRE MOF 2014 database were retained. CO<sub>2</sub> and N<sub>2</sub> adsorption isotherms were predicted using GCMC simulations, followed by process simulations under varying process variables (*e.g.*, flue gas flow rate, duration of adsorption step, and adsorption pressure). This initial

screening identified 392 MOFs that achieved the purity and recovery requirements. Further process optimization revealed that 97 MOFs outperformed amine-based capture systems, with the best MOF achieving a 25% reduction in energy consumption. Additionally, they evaluated commonly used geometric features (*e.g.*, void fraction, pore diameter, surface area) and material-level metrics (*e.g.*, CO<sub>2</sub> working capacity, Henry's selectivity, enthalpy of adsorption) in predicting process-level performance. Weak correlations between these simple metrics and process performance underscores the necessity of full process simulation and optimization for accurately evaluating MOF's performance in VSA-based carbon capture processes.

Very recently, alongside the update of the CoRE MOF database, Zhao *et al.*<sup>20</sup> selected 891 candidate materials based on criteria such as economic feasibility, structural stability, and hydrophobicity. Among these, 34 MOFs were identified by GCMC simulations to exhibit higher or similar CO<sub>2</sub>/N<sub>2</sub> selectivity and CO<sub>2</sub> working capacity than those of the benchmark material, CALF-20.<sup>14</sup> Through detailed temperature swing adsorption (TSA) simulations, the top-performing candidate was shown to outperform CALF-20 under various feed compositions and humidity levels in terms of CO<sub>2</sub> purity and recovery, highlighting its strong potential for industrial carbon capture applications.

**2.2.2. Empirical and surrogate model.** Simulation of adsorption processes can be completed in seconds to minutes, while optimization of these processes to determine their best performance may take hours. When applied to thousands of material candidates, this computational demand makes process-level screening of all materials prohibitively expensive. As discussed earlier, multi-scale screening approaches are usually used, where initial screening can be performed using various simple metrics and simplified process models, while accurate yet computationally expensive process optimization is carried out for a selected group of promising materials. To facilitate the performance-based screening of large materials databases, empirical models can be developed to correlate process-level performance with material-level properties. Furthermore, surrogate process models with high accuracy and computational efficiency can be developed as alternatives to rigorous process models,<sup>68</sup> enabling efficient large-scale materials screening with the incorporation of process-level considerations.

Developing property–performance relationships enables rapid screening of adsorbents with the consideration of their optimal performance in adsorption processes. Khurana and Farooq<sup>69</sup> proposed a method to correlate equilibrium isotherm characteristics of adsorbents with the optimal performance of the VSA process for post-combustion CO<sub>2</sub> capture. They derived mathematical equations to estimate the expected minimum energy consumption and maximum productivity of adsorbents using five characteristics related to the CO<sub>2</sub> isotherm: selectivity over N<sub>2</sub>, equilibrium adsorption capacity, local slope and nonlinearity at feed concentration,



and Henry's constant. The parameters were regressed using the results from detailed process optimization for 32 adsorbents, achieving determination coefficients of approximately 0.9 for both energy consumption and productivity. By establishing property–performance relationships, these models enable the efficient evaluation and ranking of adsorbents based on their performance in adsorption processes.

Leperi *et al.*<sup>6</sup> developed a simple evaluation metric that can effectively capture process-level information to identify the best adsorbent material with the lowest CO<sub>2</sub> capture cost. Initially, MOFs were screened based on metal costs and simple heuristics derived from preliminary process-level investigations, such as enthalpy of adsorption and adsorbent density, reducing the CoRE MOF 2014-DDEC database<sup>24</sup> to 369 candidates. For 190 MOFs that achieved the purity and recovery requirements, process simulation, optimization, and economic analysis were performed using a four-step fractionated vacuum pressure swing adsorption (FVPSA) cycle. On this basis, a general evaluation metric (GEM) was developed as a property–performance relationship using four material-level metrics: N<sub>2</sub> working capacity, CO<sub>2</sub> working capacity, CO<sub>2</sub>/N<sub>2</sub> selectivity at desorption conditions, and N<sub>2</sub> internal energy of adsorption. The GEM demonstrated superior performance in evaluating CO<sub>2</sub> capture costs, exhibiting a Spearman correlation coefficient of 0.86. Therefore, the GEM can be used to preliminarily screen adsorbent materials for post-combustion CO<sub>2</sub> capture without conducting computationally expensive process modeling and optimization.

Several research groups have demonstrated the potential of surrogate models for the efficient simulation and optimization of adsorption processes.<sup>70–77</sup> For example, Leperi *et al.*<sup>70</sup> used artificial neural networks (ANNs) to develop surrogate models for every step involved in PSA cycles (*e.g.*, pressurization, adsorption, and depressurization). Demonstrated on three different PSA cycle configurations for post-combustion CO<sub>2</sub> capture, these ANN models showed good agreement with rigorous models in process simulation while achieving a computational speedup of approximately three orders of magnitude. These results highlight the accuracy and efficiency of these ANN-based surrogate models and their potential to significantly accelerate the simulation of PSA-based carbon capture processes and performance-based materials screening.

Embedding accurate yet computationally efficient surrogate models into optimization frameworks can yield high-quality solutions at significantly reduced computational costs. However, these accurate surrogate models usually benefit from large datasets that capture full system behaviors across the design space of interest, incurring resource-intensive data collection. An alternative approach, Bayesian optimization, can be used for efficient surrogate-based optimization without relying on large datasets.<sup>78–81</sup> Using limited data, Bayesian optimization utilizes a surrogate model to iteratively propose optimal solutions for validation, with the model being continuously updated as the dataset

expands throughout the process. Recently, Bayesian optimization has been employed for the optimization of adsorption processes.<sup>82,83</sup> For example, Ward and Pini<sup>83</sup> utilized Bayesian optimization for four-step pressure-vacuum swing adsorption (PVSA) cycle in post-combustion CO<sub>2</sub> capture. Bayesian optimization was able to achieve an essentially identical solution to the conventional method that couples genetic algorithms with detailed process simulations, while reducing the computational cost by a factor of 14. This demonstrated the potential of Bayesian optimization for reliable and efficient optimization of dynamic adsorption processes in post-combustion CO<sub>2</sub> capture.

The aforementioned surrogate models are developed to estimate performance for a specific adsorbent under varying operating conditions. To make these models generally suitable for adsorbent screening, it is necessary to incorporate adsorbent features as variables. To address this, a machine-assisted adsorption process learning and emulation (MAPLE) framework<sup>84</sup> was proposed to develop generalized surrogate models that predict process performance using both adsorbent and process features (Fig. 3). Based on a wide range of inputs such as adsorbent properties, Langmuir adsorption isotherm parameters, and operating conditions, the MAPLE predicts the material's performance indicators such as product purity, recovery, energy consumption, and productivity. Specifically, using 5 isotherm parameters and 5 process variables, the MAPLE surrogates achieve an average determination coefficient exceeding 0.995 for all four performance indicators across diverse adsorbents and operating conditions. Using zeolite 13X for post-combustion CO<sub>2</sub> capture, MAPLE-based optimization achieved results nearly identical to those obtained from rigorous process optimization at significantly lower computational costs. This highlights the high accuracy and efficiency of the MAPLE framework. In another demonstration, MAPLE was applied to screen 36 adsorbents reported by Khurana and Farooq.<sup>69</sup> For 23 adsorbents that meet the CO<sub>2</sub> purity and recovery requirements, surrogate-based optimization was conducted to minimize energy consumption. The results correlated well with those from rigorous process optimization. This demonstrated the reliability of MAPLE for performance-based screening of large materials databases.

### 3. Computer-aided design of metal–organic frameworks

The chemical space of MOFs is vast and highly versatile, arising from the combination of metal nodes, organic linkers, and framework topologies. This design space is inherently complex, offering almost limitless potential structures. Consequently, extensive research efforts are dedicated to exploring this complex design space to identify novel MOF structures with superior properties for carbon capture applications.



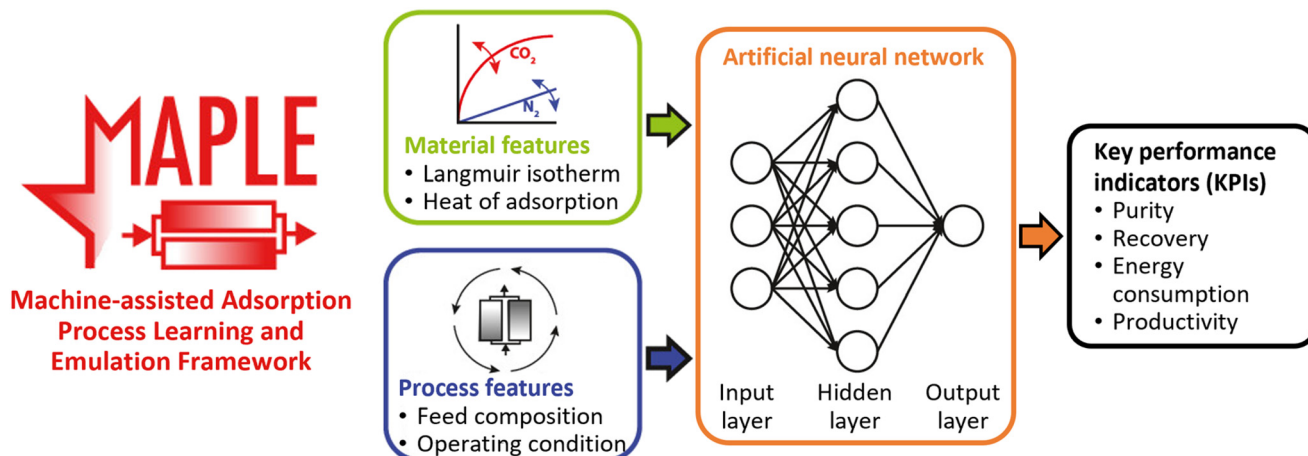


Fig. 3 The MAPLE framework for rapid simulation, optimization, and adsorbent screening of adsorption processes. By leveraging inputs such as adsorption isotherm parameters and operating conditions, the artificial neural network is used to train surrogate models to predict key performance indicators. Adapted with permission from Pai *et al.*<sup>84</sup> Copyright 2020 American Chemical Society.

This section focuses on materials design approaches for MOFs, highlighting the transition from exhaustive search in an open-loop manner to optimization-guided materials discovery in a closed-loop manner (Fig. 4). In both strategies, structure generation and evaluation are two critical components. Novel MOF structures can be generated through the functionalization of existing structures, recombination of known building blocks, or sampling from the learned latent space of generative models. Once generated, the material properties and process performance of these novel MOF structures can be evaluated using molecular simulations, structure–property relationship models, rigorous process models, and surrogate process

models, as introduced earlier. On this basis, the optimal MOF structures can be identified.

### 3.1. Open-loop design: exhaustive search

Constructed from 8 metal nodes, 94 organic linkers, and various functional groups, Boyd *et al.*<sup>42</sup> systematically generated a library of 325 000 hypothetical MOFs for capturing CO<sub>2</sub> from wet flue gases. CO<sub>2</sub>/N<sub>2</sub> selectivity and CO<sub>2</sub> working capacity of each material were evaluated using GCMC simulations, resulting in a reduced list of 8325 materials that can potentially surpass zeolite 13X under dry conditions. Based on further considerations such as

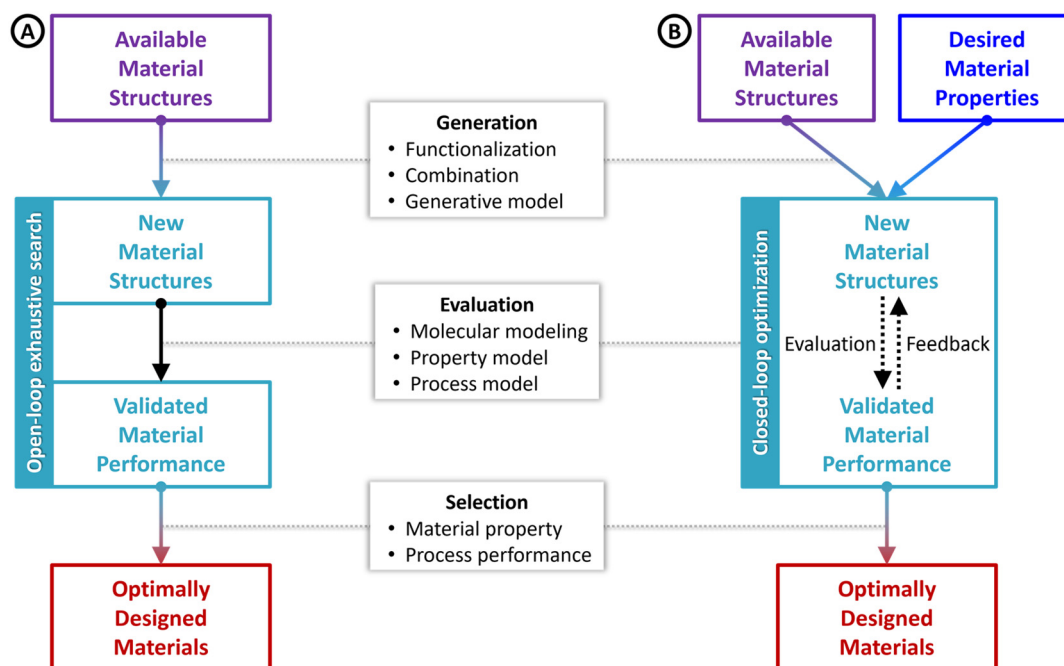


Fig. 4 General materials design workflows: (A) exhaustive search in an open-loop manner and (B) optimization-guided discovery in a closed-loop manner.



material rigidity, ease of synthesis, and hydrolytic stability of the metal–organic coordination, two hydrophobic MOFs were selected for synthesis. Experimental results confirmed that both materials outperformed commercial materials such as zeolite 13X and activated carbon, and their performance in carbon capture was not affected by water. This demonstrated that the proposed generate-and-evaluate method can effectively discover superior MOF structures for post-combustion carbon capture.

Park *et al.* proposed the GHP-MOFassemble framework to design MOF structures with high adsorption capacity and synthesizable linkers for carbon capture.<sup>85</sup> Based on organic linkers extracted from high-performing structures in the hMOF database, DiffLinker<sup>86</sup>—a diffusion model developed for molecular design—was used to generate new organic linkers. The synthesizability of these linkers was assessed using the synthetic accessibility score (SAscore)<sup>87</sup> and the synthetic complexity score (SCscore).<sup>88</sup> Synthesizable organic linkers were then assembled with pre-selected metal nodes, creating a database of 40 000 new structures. MOF screening, based on CO<sub>2</sub> adsorption capacity predicted by crystal graph convolutional neural network,<sup>89</sup> identified 364 MOF candidates. Further validation through molecular dynamics and GCMC simulations showed that 6 of these candidates exhibit CO<sub>2</sub> adsorption capacity higher than 2 mmol g<sup>-1</sup>, surpassing most structures in the hMOF database.

Leveraging generative flow networks (GFlowNets), Cipcigan *et al.* generate diverse and novel reticular materials for carbon capture applications.<sup>90</sup> Using available metal nodes and organic linkers from pormake,<sup>91</sup> they trained the GFlowNets to generate new MOF structures with improved gravimetric surface area. From 1 709 126 generated structures, those with high surface area were selected for further evaluation. GCMC simulations confirmed that all selected structures exhibited very high CO<sub>2</sub> working capacities, ranking in the 90th percentile of the CoRE MOF 2019 database. Furthermore, through structural relaxation and detailed analysis, two materials were validated to surpass all structures in the CoRE MOF 2019 database in terms of CO<sub>2</sub> working capacity. This demonstrated the potential of GFlowNets as a powerful tool for designing promising MOFs for carbon capture applications.

### 3.2. Closed-loop design: optimization-guided discovery

The generation and evaluation of material structures are essential for discovering new materials with improved properties. However, the performance of optimal candidates is often constrained due to the limited exploration of the design space. Additionally, in the open-loop design, the desired properties cannot be directly targeted during the structure generation. To address these challenges, results from the evaluation phase should be fed back into the generation phase, closing the loop of materials design. By incorporating optimization algorithms, the generation of materials can be guided to satisfy desired properties,

enabling property-oriented materials design with significantly reduced computational costs. This closed-loop, optimization-guided approach allows for efficient exploration of large design spaces compared to exhaustive search methods, discovering new MOF structures with optimized properties. This approach is also referred to as inverse materials design, which begins with the desired functionality and searches for ideal material structures within the design space using optimization techniques.<sup>95</sup>

Evolutionary algorithms, such as genetic algorithms (GAs), can be used to iteratively optimize material structures and improve their properties. Chung *et al.*<sup>35</sup> developed a GA to efficiently identify top-performing MOFs for precombustion CO<sub>2</sub> capture from thousands of candidates. Each MOF structure was represented as a sequence of integer chemical identities, and the integer combination was optimized by the GA to improve material properties. Using a database of 55 163 hypothetical MOFs, the optimization-guided search was conducted with three different objectives: CO<sub>2</sub> working capacity, CO<sub>2</sub>/H<sub>2</sub> selectivity, and APS. Compared to an exhaustive search, the optimization-guided approach significantly reduced the number of required GCMC simulations, as only 730 MOFs were evaluated across all three tasks. Experimental synthesis and analysis of the best MOF confirmed its high CO<sub>2</sub> working capacity and CO<sub>2</sub>/H<sub>2</sub> selectivity, demonstrating the effectiveness of GAs in optimization-guided MOF discovery.

Deng and Sarkisov<sup>96</sup> developed a computational design workflow using GA to optimize MOF structures for post-combustion CO<sub>2</sub> capture. Starting with a library of building blocks and topologies, they generated tens of thousands of MOF structures using pormake.<sup>91</sup> The GA then optimized the combination of building blocks and topologies to design new structures, while MOF-NET<sup>91</sup> efficiently predicted N<sub>2</sub> and CO<sub>2</sub> adsorption capacities. To validate the performance of the optimized structures, high-fidelity GCMC simulations and process optimization were conducted in a modified Skarstrom cycle. Most of the designed MOFs presented significantly better performance than zeolite 13X in terms of energy consumption and productivity, demonstrating the effectiveness of GA in MOF design.

Functionalization is a key strategy to improve the functional properties of MOFs. Collins *et al.*<sup>97</sup> developed MOFF-GA—a customized genetic algorithm (GA)—to optimize functional groups in MOF structures. Validated on a diverse set of 48 MOFs, MOFF-GA effectively identified the best functionalized structures while exploring only a small fraction of the vast search space. During the functionalization process, geometry optimization was performed for newly generated structures, and their CO<sub>2</sub> adsorption properties were evaluated using GCMC simulations. When applied to optimize the CO<sub>2</sub> adsorption capacity of 141 experimentally characterized MOFs for post-combustion CO<sub>2</sub> capture, MOFF-GA identified 1035 functionalized structures with exceptional capacities exceeding 3 mmol g<sup>-1</sup> at 0.15 atm and 298 K. These results highlight MOFF-GA as an efficient approach for



designing high-performance MOFs for carbon capture through functionalization.

Zhang *et al.*<sup>98</sup> developed a computational framework to design MOFs tailored for carbon capture. Specifically, organic linkers were first generated using the Monte Carlo tree search (MCTS) implemented in ChemTS,<sup>99</sup> and then they were combined with 10 different metal node and topology combinations to construct new MOF structures. The adsorption properties of each generated structure were evaluated using GCMC simulations, with the results fed back into the MCTS algorithm to refine the structure generation. This approach was demonstrated very efficiently, as the design and GCMC validation of five MOFs can be completed within one hour. On average, the optimal MOF was identified within 86.1 MCTS cycles, achieving a 134.63% improvement in CO<sub>2</sub> adsorption capacity compared to the original MOF. These findings highlight the effectiveness of the proposed approach in designing MOFs with significantly improved properties. In a recent study, the researchers successfully extended their approach to design MOFs for efficient carbon capture under humid conditions.<sup>100</sup>

Reinforcement learning (RL) is a machine learning technique that learns to make decisions to achieve the most optimal outcomes. Park *et al.* introduced a deep reinforcement learning approach for designing MOFs with desired properties.<sup>101</sup> The RL framework consists of two key components: an agent for decision-making (*i.e.*, structure generation) and an environment for evaluation (*i.e.*, property

estimation). The agent assembles MOF structures by selecting metal clusters, organic linkers, and topologies, while the environment evaluates their properties and returns rewards of property improvement to the agent. This iterative process refines structure generation until the agent designs MOFs with the desired properties. Applied to carbon capture, the RL framework successfully designed a set of MOFs that have high CO<sub>2</sub> enthalpy of adsorption and can selectively adsorb CO<sub>2</sub> from humid air. Notably, some of these MOFs exhibited higher calculated CO<sub>2</sub> enthalpy of adsorption and CO<sub>2</sub>/H<sub>2</sub>O selectivity than previously reported MOFs. This demonstrated the potential of RL as a powerful tool for designing MOFs with improved targeted properties.

Generative models, such as variational autoencoders (VAEs)<sup>102</sup> and generative adversarial networks (GANs),<sup>103</sup> can be used as key components in enabling closed-loop materials design for targeted properties. Leveraging a VAE, Yao *et al.* developed an automated discovery platform for property-oriented MOF design.<sup>102</sup> MOF structures were represented by edges, vertices, and topologies decomposed from reticular frameworks, which in turn enables their reconstruction (Fig. 5). Using 372 edges identified from the CoRE MOF 2019 database, an augmented dataset of around 300 000 edges was generated through functionalization. Combining this edge dataset with vertex and topology datasets, a dataset containing approximately two million MOF structures was constructed. To enable property-oriented design, computational simulations were performed on around 45 000

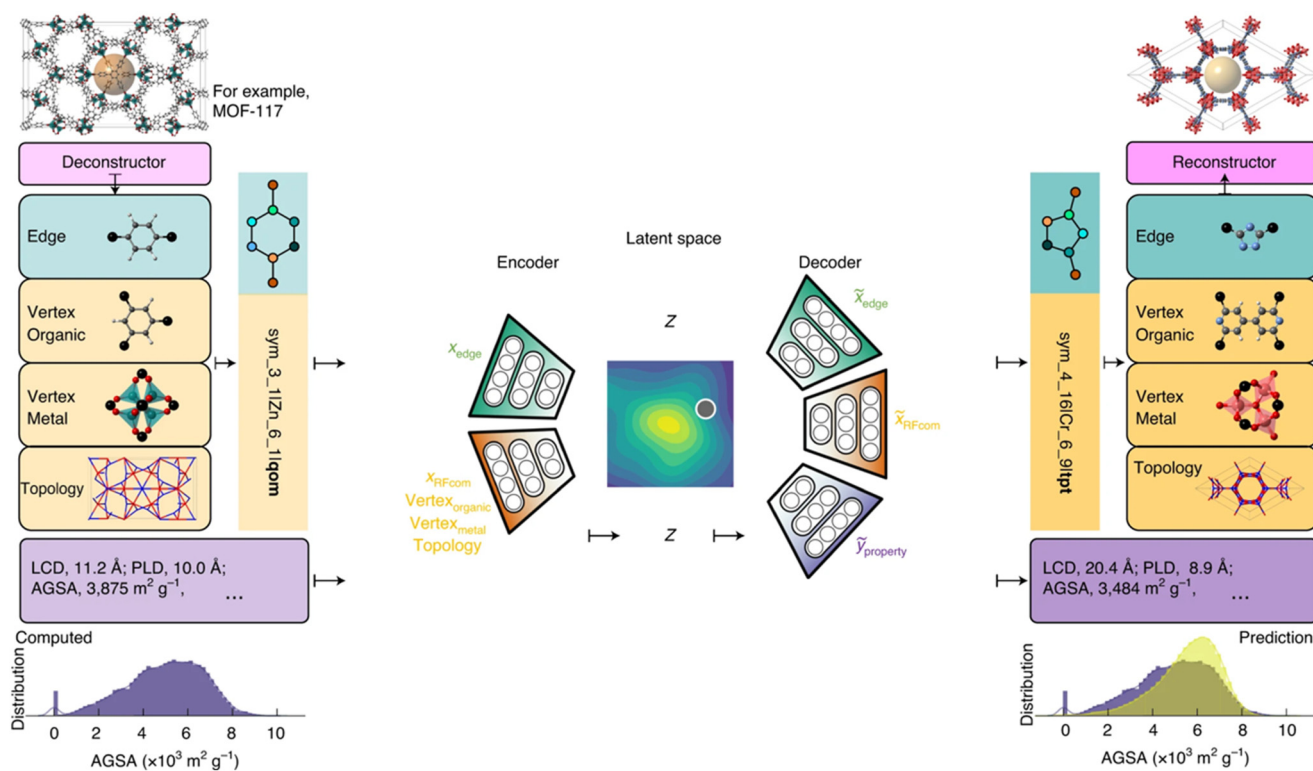


Fig. 5 Schematic of the automated materials design platform empowered by supramolecular variational autoencoder. Reproduced with permission from Yao *et al.*,<sup>102</sup> Springer Nature.



randomly selected MOFs to train the VAE, alongside the augmented structure dataset. The trained model was then used to generate and optimize MOF structures with improved properties for CO<sub>2</sub> separation from natural gas. Confirmed *via* GCMC simulations, the generated top-performing MOF achieved a CO<sub>2</sub> adsorption capacity of 7.55 mol kg<sup>-1</sup> and a CO<sub>2</sub>/CH<sub>4</sub> selectivity of 16. Overall, this generative model enables optimization-guided materials discovery, accelerating the design of materials with tailored adsorption properties for carbon capture.

Very recently, building on the “deep dreaming” methodology implemented in inverse molecular design,<sup>92</sup> Cleeton *et al.*<sup>93</sup> extended the concept to the inverse design of MOFs. In their design framework, MOF structures are represented using SELFIES string representations<sup>94</sup> due to their guaranteed chemical validity when decoded into molecular structures. A machine learning model based on long short-term memory networks was trained to predict material properties. The model architecture was then inverted for deep dreaming to design new MOFs with improved property values. Specifically, the model weights and biases were frozen, and the input representations were iteratively optimized toward desired properties, such as specific heat capacity and CO<sub>2</sub>/N<sub>2</sub> Henry selectivity. For MOF design, deep dreaming achieves high validity (~95.5%) and uniqueness (~99.4%) of the generated structures. However, its design scope is narrow since the optimization is limited to the linkers only. Despite this limitation, deep dreaming represents one of several emerging strategies in inverse design and has shown promise for designing new MOFs with targeted properties relevant to carbon capture applications.

## 4. Conclusions and perspectives

Significant progress has been achieved in recent years in discovering promising MOFs for diverse carbon capture applications through various computer-aided screening and design approaches. These include property- and performance-based materials screening, as well as materials design in both open-loop and closed-loop manners. However, challenges remain in improving the reliability and efficiency of these computational approaches. In the following, we discuss possible solutions to these challenges and outline future research directions to advance computer-aided materials screening and design, aimed at discovering high-performance MOFs for practical carbon capture applications.

### Accuracy of molecular simulations

Accurately describing interatomic interactions is essential for reliable molecular simulations. In classical GCMC simulations, this is typically achieved using a force field, *i.e.*, a set of equations and parameters that describe both short-range and long-range interactions. Short-range repulsion and dispersion interactions are described using Lennard-Jones (LJ) potentials, while long-range electrostatic interactions are described using Coulomb potentials.<sup>104</sup>

Consequently, the accuracy of GCMC simulations strongly depends on the choice of force field, particularly the LJ parameters and partial atomic charges. In computational studies, different combinations of LJ parameters and charge assignment schemes are employed, depending on factors such as feed composition, material characteristics, and resource availability.

A recent systematic study by Cleeton *et al.*<sup>104</sup> demonstrated that both LJ parameters and partial atomic charges influence the reliability of simulation results and, therefore, the ranking of MOFs in PSA-based carbon capture processes. Notably, the choice of charge assignment scheme was found to have a more decisive impact. Based on these findings, the authors advocate for using partial charges derived from *ab initio* calculations, particularly when electrostatic interactions dominate (such as CO<sub>2</sub> and H<sub>2</sub>O adsorption in MOFs). Moreover, GCMC simulations of H<sub>2</sub>O adsorption are more approximate and challenging due to the complex nature of water-MOF interactions.<sup>105</sup> In such cases, employing high-fidelity charge assignment schemes becomes even more critical to accurately capture adsorption behaviors. When *ab initio* charge calculations are computationally prohibitive, machine learning methods offer an attractive alternative. Recent ML models, such as message passing neural network,<sup>106</sup> PACMOF,<sup>107</sup> PACMAN,<sup>108</sup> and MEPO-ML,<sup>109</sup> have shown the ability to predict partial charges with accuracy comparable to *ab initio* methods. These ML-based approaches enable efficient yet accurate MOF screening and design for carbon capture, where electrostatic interactions play a critical role in adsorption behavior.

### Stability, synthesizability, and commercialization

Although many MOF candidates demonstrate exceptional performance in molecular simulations and experimental validations, maintaining their stability and functionality under industrial conditions remains a significant challenge. This is particularly true for carbon capture applications, where MOF structures are often prone to degradation under harsh environments such as high temperatures, humidity, and acidic gases. The evaluation of MOF stability primarily relies on experimental techniques such as thermogravimetric analysis (TGA) and powder X-ray diffraction (PXRD). However, these methods are not easily scalable for evaluating the large number of MOFs. To address this, researchers have leveraged reported experimental stability data to train machine learning models that capture structure–stability relationships, enabling the prediction of diverse stabilities such as thermal, solvent removal, and water stabilities.<sup>110–114</sup> In addition to stability, synthesizability is also a major bottleneck. Many high-performing MOFs are difficult to synthesize because of diverse challenges such as difficulties in synthesizing organic linkers.<sup>115</sup> Systematically evaluating MOF synthesizability is still an open and important problem. Although several metrics such as synthetic accessibility (SA),<sup>87</sup> synthetic Bayesian accessibility (SYBA),<sup>116</sup> synthetic complexity score



(SCScore),<sup>88</sup> and retrosynthetic accessibility score (RAScore)<sup>117</sup> are available to estimate the ease of synthesizing organic molecules, their applicability is limited to full MOF structures. As a result, the assessment of MOF synthesizability still relies on expert intuition, which limits the scalability of computational MOF screening and design. To mitigate this, a computational tool called MOFSynth,<sup>118</sup> was recently developed to predict the likelihood of MOF synthesizability, serving as an initial filter to eliminate candidates with poor synthesizability. Thus, incorporating both stability and synthesizability considerations is essential for accelerating the discovery of promising, stable, and synthesizable MOFs for industrial-scale deployment of carbon capture technologies.

Despite the promise of MOFs in carbon capture, only a handful of MOFs have successfully transitioned from fundamental research to commercialization. To successfully realize this transition, five key topics must be addressed in applied research, including synthesis, forming, processing (washing and activation), prototyping, and compliance.<sup>119</sup> Successfully overcoming these hurdles requires considerable resources, a multidisciplinary team, and close collaboration between academic institutions and industry partners.

### High-quality experimental databases

Computational development of reliable materials for carbon capture can significantly benefit from experimental insights. However, collecting high-fidelity experimental data remains challenging and resource intensive. Over the past decade, the NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials<sup>120</sup> has been developed to include adsorption experiments and isotherms for various gases (both pure and mixture) and adsorbent materials, digitized from the literature. Additionally, related research has improved the querying and analysis of material properties within this database.<sup>121,122</sup> Despite these significant efforts, issues regarding data reusability still exist. For example, adsorption and desorption measurements were sometimes mis-integrated into a single isotherm, and adsorption data for pure components and mixtures was occasionally mislabeled. In such cases, problematic experimental data can significantly affect the reliability of computer-aided discovery of adsorbent materials. To address this, numerical data underlying key figures should be provided along with the publication to enhance reproducibility and data reusability. This can help us integrate large synthetic datasets with small experimental datasets to facilitate reliable materials discovery.

Several studies have used data mining to collect experimental data from the literature.<sup>111,123,124</sup> Natural language processing techniques were typically used to automatically analyze and extract experimental data, enabling the efficient digitization of large volumes of information into databases. However, challenges of data reusability arose during such digitalization processes. Due to challenges such as diverse data types, a small fraction of the extracted data

was incorrect, potentially misleading the downstream materials development and analysis. Therefore, the establishment of high-quality experimental databases still necessitates human-machine collaboration to ensure the accuracy and reliability of the digitalization process.

On the other hand, the structural fidelity of MOF databases remains a critical issue for reliable materials screening. Incorrect assignment of metal oxidation states can significantly affect a MOF's electronic structure, leading to drastically deviate properties in molecular simulation.<sup>125</sup> For instance, mediocre MOFs may be mistakenly identified as top candidates. To address this problem, computational tools such as MOSAEC<sup>125</sup> and MOFChecker<sup>126</sup> have been developed to automatically detect structural errors and filter out problematic structures. By incorporating such validation tools, the reliability of MOF candidates identified through computational screening and design can be significantly improved.

### Active materials discovery

Molecular simulation can exhaustively identify optimal candidates, whereas it is computationally prohibitive for large materials databases. Therefore, adaptive strategies are essential to efficiently search for optimal materials with minimal resource consumption. A possible approach is Bayesian optimization, which integrates machine learning, uncertainty quantification, and informed decision-making to automatically and efficiently identify optimal materials.<sup>127,128</sup> By implementing Bayesian optimization, an iterative process consisting of molecular simulation, surrogate modeling of the simulations, and selection of candidates for evaluation is conducted until a promising material is identified. In this way, the computational costs are significantly reduced as only the most promising candidates identified by Bayesian optimization are evaluated using time-consuming molecular simulations.

Structure–property relationship models have significantly accelerated property-based materials screening. However, the accuracy of such models heavily relies on the availability of data, which is resource-intensive to collect. In this context, adaptive strategies such as active learning can help to develop accurate models with minimal data demand.<sup>78</sup> Taking advantage of active learning, the model development requires fewer molecular simulations to obtain an accurate model covering the material space of interest. This can significantly reduce the costs associated with data generation while maintaining the high accuracy and applicability of models for subsequent materials screening. These adaptive strategies enable active materials discovery, showing great potential to accelerate the development of high-performance materials for adsorption-based carbon capture applications.

### Integrated material and process design

For performance-based large-scale materials screening, candidates that survive property-based preselection (*i.e.*, structure–property relationships) are further evaluated with process simulation and optimization (*i.e.*, property–



performance relationships) to determine their optimal performance in carbon capture processes. To account for the interplay between materials and varying operating conditions, process simulation and optimization should be directly integrated with the selection of materials. This allows for the simultaneous identification of the optimal material and its tailored operating conditions for practical adsorption processes.<sup>129,130</sup> By considering the entire structure–property–performance hierarchy, this approach enables process-informed materials discovery through integrated material and process design.

Current closed-loop materials design workflows primarily target specific material properties. By incorporating process simulation and optimization, the focus can shift from property-oriented to performance-oriented design, ensuring that materials are optimally designed with actual process performance in mind. Diverse key performance indicators (KPIs) on the process level—such as product purity and recovery, energy consumption, productivity, capital and operating costs, and environmental impact—can be considered in evaluating carbon capture systems. Trade-offs between these often-competing objectives can be identified through multi-objective optimization and can be represented as a set of Pareto-optimal solutions, offering design schemes for diverse industrial carbon capture scenarios. Furthermore, as implemented in the PrISMa (process-informed design of tailor-made sorbent materials) platform,<sup>131</sup> techno-economic analysis and life-cycle assessment can be integrated to provide a holistic evaluation of material performance, thereby accelerating the development and deployment of carbon capture technologies toward a net-zero future. This performance-oriented closed-loop materials design can be further accelerated by data-driven models, which rapidly assess the feasibility and efficiency of separation systems by incorporating considerations from both material and process scales.<sup>132,133</sup> Ultimately, this approach enables a fully integrated material and process design in a closed-loop manner, advancing the discovery of high-performance adsorbent materials for carbon capture applications.

## Author contributions

Zihao Wang: conceptualization; visualization; writing – original draft; writing – review & editing. Teng Zhou: conceptualization; supervision; funding acquisition; writing – review & editing.

## Conflicts of interest

There are no conflicts of interest to declare.

## Data availability

No primary research results, software or code have been included and no new data were generated or analysed as part of this review.

## Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 22308290), Guangdong Provincial Project (2023QN10L094), Guangzhou Basic and Applied Basic Research Program (2025A04J7024), and Nansha Key Science and Technology Project (2023ZD013). Teng Zhou additionally acknowledges financial support from the Max Planck Society through the Max Planck Partner Group Program.

## References

- 1 M. Bui, C. S. Adjiman and A. Bardow, *et al.*, Carbon capture and storage (CCS): the way forward, *Energy Environ. Sci.*, 2018, **11**(5), 1062–1176.
- 2 A. H. Farmahini, S. Krishnamurthy, D. Friedrich, S. Brandani and L. Sarkisov, Performance-based screening of porous materials for carbon capture, *Chem. Rev.*, 2021, **121**(17), 10666–10741.
- 3 J. M. Huck, L.-C. Lin and A. H. Berger, *et al.*, Evaluating different classes of porous materials for carbon capture, *Energy Environ. Sci.*, 2014, **7**(12), 4132–4146.
- 4 Y. Hu, R. Gani, K. Sundmacher and T. Zhou, Assessing the future impact of 12 direct air capture technologies, *Chem. Eng. Sci.*, 2024, **298**, 120423.
- 5 X. Ding, H. Chen, J. Li and T. Zhou, Comparative techno-economic analysis of CO<sub>2</sub> capture processes using blended amines, *Carbon Capture Sci. Technol.*, 2023, **9**, 100136.
- 6 K. T. Leperi, Y. G. Chung, F. You and R. Q. Snurr, Development of a general evaluation metric for rapid screening of adsorbent materials for postcombustion CO<sub>2</sub> capture, *ACS Sustainable Chem. Eng.*, 2019, **7**(13), 11529–11539.
- 7 L.-C. Lin, A. H. Berger and R. L. Martin, *et al.*, In silico screening of carbon-capture materials, *Nat. Mater.*, 2012, **11**(7), 633–641.
- 8 C. A. Grande, Advances in pressure swing adsorption for gas separation, *Int. Scholarly Res. Not.*, 2012, **2012**(1), 982934.
- 9 K. Sumida, D. L. Rogow and J. A. Mason, *et al.*, Carbon dioxide capture in metal–organic frameworks, *Chem. Rev.*, 2012, **112**(2), 724–781.
- 10 M. F. Hasan, E. L. First and C. A. Floudas, Cost-effective CO<sub>2</sub> capture based on in silico screening of zeolites and process optimization, *Phys. Chem. Chem. Phys.*, 2013, **15**(40), 17601–17618.
- 11 M. S. Khosrowshahi, A. A. Aghajari and M. Rahimi, *et al.*, Recent progress on advanced solid adsorbents for CO<sub>2</sub> capture: from mechanism to machine learning, *Mater. Today Sustain.*, 2024, 100900.
- 12 J.-R. Li, Y. Ma and M. C. McCarthy, *et al.*, Carbon dioxide capture-related gas adsorption and separation in metal–organic frameworks, *Coord. Chem. Rev.*, 2011, **255**(15–16), 1791–1823.
- 13 R.-B. Lin, S. Xiang, H. Xing, W. Zhou and B. Chen, Exploration of porous metal–organic frameworks for gas



- separation and purification, *Coord. Chem. Rev.*, 2019, **378**, 87–103.
- 14 J.-B. Lin, T. T. Nguyen and R. Vaidhyanathan, *et al.*, A scalable metal-organic framework as a durable physisorbent for carbon dioxide capture, *Science*, 2021, **374**(6574), 1464–1469.
  - 15 C. Zhang, Y. Xie, C. Xie, H. Dong, L. Zhang and J. Lin, Accelerated discovery of porous materials for carbon capture by machine learning: A review, *MRS Bull.*, 2022, **47**(4), 432–439.
  - 16 H. Daglar and S. Keskin, Recent advances, opportunities, and challenges in high-throughput computational screening of MOFs for gas separations, *Coord. Chem. Rev.*, 2020, **422**, 213470.
  - 17 C. E. Wilmer, M. Leaf and C. Y. Lee, *et al.*, Large-scale screening of hypothetical metal-organic frameworks, *Nat. Chem.*, 2012, **4**(2), 83–89.
  - 18 Y. G. Chung, J. Camp and M. Haranczyk, *et al.*, Computation-ready, experimental metal-organic frameworks: A tool to enable high-throughput screening of nanoporous crystals, *Chem. Mater.*, 2014, **26**(21), 6185–6192.
  - 19 Y. G. Chung, E. Haldoupis and B. J. Bucior, *et al.*, Advances, updates, and analytics for the computation-ready, experimental metal-organic framework database: CoRE MOF 2019, *J. Chem. Eng. Data*, 2019, **64**(12), 5985–5998.
  - 20 G. Zhao, L. M. Brabson and S. Chheda, *et al.*, CoRE MOF DB: A curated experimental metal-organic framework database with machine-learned properties for integrated material-process screening, *Matter*, 2025, **8**(6), 102140.
  - 21 N. S. Bobbitt, K. Shi and B. J. Bucior, *et al.*, MOFX-DB: An online database of computational adsorption data for nanoporous materials, *J. Chem. Eng. Data*, 2023, **68**(2), 483–498.
  - 22 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, The Cambridge structural database, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 2016, **72**(2), 171–179.
  - 23 P. G. Boyd and T. K. Woo, A generalized method for constructing hypothetical nanoporous materials of any net topology from graph theory, *CrystEngComm*, 2016, **18**(21), 3777–3792.
  - 24 D. Nazarian, J. S. Camp and D. S. Sholl, A comprehensive set of high-quality point charges for simulations of metal-organic frameworks, *Chem. Mater.*, 2016, **28**(3), 785–793.
  - 25 Y. J. Colón, D. A. Gómez-Gualdrón and R. Q. Snurr, Topologically guided, automated construction of metal-organic frameworks and their evaluation for energy-related applications, *Cryst. Growth Des.*, 2017, **17**(11), 5801–5810.
  - 26 A. S. Rosen, S. M. Iyer and D. Ray, *et al.*, Machine learning the quantum-chemical properties of metal-organic frameworks for accelerated materials discovery, *Matter*, 2021, **4**(5), 1578–1597.
  - 27 A. Nandy, S. Yue and C. Oh, *et al.*, A database of ultrastable MOFs reassembled from stable fragments with machine learning models, *Matter*, 2023, **6**(5), 1585–1603.
  - 28 J. Burner, J. Luo and A. White, *et al.*, ARC-MOF: a diverse database of metal-organic frameworks with DFT-derived partial atomic charges and descriptors for machine learning, *Chem. Mater.*, 2023, **35**(3), 900–916.
  - 29 T. F. Willems, C. H. Rycroft, M. Kazi, J. C. Meza and M. Haranczyk, Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials, *Microporous Mesoporous Mater.*, 2012, **149**(1), 134–141.
  - 30 L. Sarkisov, R. Bueno-Perez, M. Sutharson and D. Fairen-Jimenez, Materials informatics with PoreBlazer v4.0 and the CSD MOF database, *Chem. Mater.*, 2020, **32**(23), 9849–9867.
  - 31 A. Rajendran, S. G. Subraveti, K. N. Pai, V. Prasad and Z. Li, How can (or why should) process engineering aid the screening and discovery of solid sorbents for CO<sub>2</sub> capture?, *Acc. Chem. Res.*, 2023, **56**(17), 2354–2365.
  - 32 Y. S. Bae and R. Q. Snurr, Development and evaluation of porous materials for carbon dioxide separation and capture, *Angew. Chem., Int. Ed.*, 2011, **50**(49), 11586–11596.
  - 33 G. D. Pirngruber, L. Hamon and S. Bourrelly, *et al.*, A method for screening the potential of MOFs as CO<sub>2</sub> adsorbents in pressure swing adsorption processes, *ChemSusChem*, 2012, **5**(4), 762–776.
  - 34 A. D. Wiersum, J.-S. Chang, C. Serre and P. L. Llewellyn, An adsorbent performance indicator as a first step evaluation of novel sorbents for gas separations: application to metal-organic frameworks, *Langmuir*, 2013, **29**(10), 3301–3309.
  - 35 Y. G. Chung, D. A. Gómez-Gualdrón and P. Li, *et al.*, In silico discovery of metal-organic frameworks for precombustion CO<sub>2</sub> capture using a genetic algorithm, *Sci. Adv.*, 2016, **2**(10), e1600909.
  - 36 H. Demir, H. Daglar, H. C. Gulbalkan, G. O. Aksu and S. Keskin, Recent advances in computational modeling of MOFs: From molecular simulations to machine learning, *Coord. Chem. Rev.*, 2023, **484**, 215112.
  - 37 D. Dubbeldam, S. Calero, D. E. Ellis and R. Q. Snurr, RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials, *Mol. Simul.*, 2016, **42**(2), 81–101.
  - 38 G. Avci, S. Velioglu and S. Keskin, High-throughput screening of MOF adsorbents and membranes for H<sub>2</sub> purification and CO<sub>2</sub> capture, *ACS Appl. Mater. Interfaces*, 2018, **10**(39), 33693–33706.
  - 39 G. Avci, I. Erucar and S. Keskin, Do new MOFs perform better for CO<sub>2</sub> capture and H<sub>2</sub> purification? Computational screening of the updated MOF database, *ACS Appl. Mater. Interfaces*, 2020, **12**(37), 41567–41579.
  - 40 Z. Qiao, K. Zhang and J. Jiang, In silico screening of 4764 computation-ready, experimental metal-organic frameworks for CO<sub>2</sub> separation, *J. Mater. Chem. A*, 2016, **4**(6), 2105–2114.
  - 41 H. Xu, L. L. Mguni and Y. Shan, *et al.*, Computational prediction of MOFs with the potential to improve the efficiency of industrial CO<sub>2</sub> capture, *Sep. Purif. Technol.*, 2025, **354**, 128927.
  - 42 P. G. Boyd, A. Chidambaram and E. García-Díez, *et al.*, Data-driven design of metal-organic frameworks for wet flue gas CO<sub>2</sub> capture, *Nature*, 2019, **576**(7786), 253–256.



- 43 S. Kancharlapalli and R. Q. Snurr, High-throughput screening of the CoRE-MOF-2019 database for CO<sub>2</sub> capture from wet flue gas: a multi-scale modeling strategy, *ACS Appl. Mater. Interfaces*, 2023, **15**(23), 28084–28092.
- 44 J. Rogacka, A. Seremak and A. Luna-Triguero, *et al.*, High-throughput screening of metal–organic frameworks for CO<sub>2</sub> and CH<sub>4</sub> separation in the presence of water, *Chem. Eng. J.*, 2021, **403**, 126392.
- 45 S. Li, Y. G. Chung and R. Q. Snurr, High-throughput screening of metal–organic frameworks for CO<sub>2</sub> capture in the presence of water, *Langmuir*, 2016, **32**(40), 10368–10376.
- 46 J. Park, H. Kim, Y. Kang, Y. Lim and J. Kim, From data to discovery: recent trends of machine learning in metal–organic frameworks, *JACS Au.*, 2024, **4**(10), 3727–3743.
- 47 H. C. Gulbalkan, G. O. Aksu, G. Ercakir and S. Keskin, Accelerated discovery of metal–organic frameworks for CO<sub>2</sub> capture by artificial intelligence, *Ind. Eng. Chem. Res.*, 2023, **63**(1), 37–48.
- 48 S. Chong, S. Lee, B. Kim and J. Kim, Applications of machine learning in metal-organic frameworks, *Coord. Chem. Rev.*, 2020, **423**, 213487.
- 49 H. Dureckova, M. Krykunov, M. Z. Aghaji and T. K. Woo, Robust machine learning models for predicting high CO<sub>2</sub> working capacity and CO<sub>2</sub>/H<sub>2</sub> selectivity of gas adsorption in metal organic frameworks for precombustion carbon capture, *J. Phys. Chem. C*, 2019, **123**(7), 4133–4139.
- 50 M. Pardakhti, E. Moharrerri, D. Wanik, S. L. Suib and R. Srivastava, Machine learning using combined structural and chemical descriptors for prediction of methane adsorption performance of metal organic frameworks (MOFs), *ACS Comb. Sci.*, 2017, **19**(10), 640–645.
- 51 G. S. Fanourgakis, K. Gkagkas, E. Tylianakis and G. E. Froudakis, A universal machine learning algorithm for large-scale screening of materials, *J. Am. Chem. Soc.*, 2020, **142**(8), 3814–3822.
- 52 B. J. Bucior, N. S. Bobbitt and T. Islamoglu, *et al.*, Energy-based descriptors to rapidly predict hydrogen storage in metal–organic frameworks, *Mol. Syst. Des. Eng.*, 2019, **4**(1), 162–174.
- 53 Z. Deng and L. Sarkisov, Engineering machine learning features to predict adsorption of carbon dioxide and nitrogen in metal–organic frameworks, *J. Phys. Chem. C*, 2024, **128**(24), 10202–10215.
- 54 Z. Wang, Y. Zhou, T. Zhou and K. Sundmacher, Identification of optimal metal-organic frameworks by machine learning: Structure decomposition, feature integration, and predictive modeling, *Comput. Chem. Eng.*, 2022, **160**, 107739.
- 55 T. Xie and J. C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, *Phys. Rev. Lett.*, 2018, **120**(14), 145301.
- 56 R. Wang, Y. Zou, C. Zhang, X. Wang, M. Yang and D. Xu, Combining crystal graphs and domain knowledge in machine learning to predict metal-organic frameworks performance in methane adsorption, *Microporous Mesoporous Mater.*, 2022, **331**, 111666.
- 57 Y. Kang, H. Park, B. Smit and J. Kim, A multi-modal pre-training transformer for universal transfer learning in metal–organic frameworks, *Nat. Mach. Intell.*, 2023, **5**(3), 309–318.
- 58 M. Fernandez and A. S. Barnard, Geometrical properties can predict CO<sub>2</sub> and N<sub>2</sub> adsorption performance of metal–organic frameworks (MOFs) at low pressure, *ACS Comb. Sci.*, 2016, **18**(5), 243–252.
- 59 M. Fernandez, N. R. Trefiak and T. K. Woo, Atomic property weighted radial distribution functions descriptors of metal–organic frameworks for the prediction of gas uptake capacity, *J. Phys. Chem. C*, 2013, **117**(27), 14095–14105.
- 60 K. Choudhary, T. Yildirim, D. W. Siderius, A. G. Kusne, A. McDannald and D. L. Ortiz-Montalvo, Graph neural network predictions of metal organic framework CO<sub>2</sub> adsorption properties, *Comput. Mater. Sci.*, 2022, **210**, 111388.
- 61 I. B. Orhan, T. C. Le, R. Babarao and A. W. Thornton, Accelerating the prediction of CO<sub>2</sub> capture at low partial pressures in metal-organic frameworks using new machine learning descriptors, *Commun. Chem.*, 2023, **6**(1), 214.
- 62 Y. Teng and G. Shan, Interpretable machine learning for materials discovery: Predicting CO<sub>2</sub> adsorption properties of metal–organic frameworks, *APL Mater.*, 2024, **12**(8), 081115.
- 63 A. K. Rajagopalan, A. M. Avila and A. Rajendran, Do adsorbent screening metrics predict process performance? A process optimisation based study for post-combustion capture of CO<sub>2</sub>, *Int. J. Greenhouse Gas Control*, 2016, **46**, 76–85.
- 64 D. Yancy-Caballero, K. T. Leperi and B. J. Bucior, *et al.*, Process-level modelling and optimization to evaluate metal–organic frameworks for post-combustion capture of CO<sub>2</sub>, *Mol. Syst. Des. Eng.*, 2020, **5**(7), 1205–1218.
- 65 K. N. Pai, J. D. Baboolal, D. A. Sharp and A. Rajendran, Evaluation of diamine-appended metal-organic frameworks for post-combustion CO<sub>2</sub> capture by vacuum swing adsorption, *Sep. Purif. Technol.*, 2019, **211**, 540–550.
- 66 J. Park, H. c. O. Rubiera Landa, Y. Kawajiri, M. J. Realff, R. P. Lively and D. S. Sholl, How well do approximate models of adsorption-based CO<sub>2</sub> capture processes predict results of detailed process models?, *Ind. Eng. Chem. Res.*, 2019, **59**(15), 7097–7108.
- 67 T. D. Burns, K. N. Pai and S. G. Subraveti, *et al.*, Prediction of MOF performance in vacuum swing adsorption systems for postcombustion CO<sub>2</sub> capture based on integrated molecular simulations, process optimizations, and machine learning models, *Environ. Sci. Technol.*, 2020, **54**(7), 4536–4544.
- 68 K. McBride and K. Sundmacher, Overview of surrogate modeling in chemical process engineering, *Chem. Ing. Tech.*, 2019, **91**(3), 228–239.
- 69 M. Khurana and S. Farooq, Adsorbent screening for postcombustion CO<sub>2</sub> capture: a method relating equilibrium isotherm characteristics to an optimum



- vacuum swing adsorption process performance, *Ind. Eng. Chem. Res.*, 2016, **55**(8), 2447–2460.
- 70 K. T. Leperi, D. Yancy-Caballero, R. Q. Snurr and F. You, 110th anniversary: surrogate models based on artificial neural networks to simulate and optimize pressure swing adsorption cycles for CO<sub>2</sub> capture, *Ind. Eng. Chem. Res.*, 2019, **58**(39), 18241–18252.
- 71 A. Agarwal, L. T. Biegler and S. E. Zitney, Simulation and optimization of pressure swing adsorption systems using reduced-order modeling, *Ind. Eng. Chem. Res.*, 2009, **48**(5), 2327–2343.
- 72 M. F. Hasan, R. C. Baliban, J. A. Elia and C. A. Floudas, Modeling, simulation, and optimization of postcombustion CO<sub>2</sub> capture for variable feed concentration and flow rate. 2. Pressure swing adsorption and vacuum swing adsorption processes, *Ind. Eng. Chem. Res.*, 2012, **51**(48), 15665–15682.
- 73 F. Boukouvala, M. F. Hasan and C. A. Floudas, Global optimization of general constrained grey-box models: new method and its application to constrained PDEs for pressure swing adsorption, *J. Glob. Optim.*, 2017, **67**, 3–42.
- 74 S. G. Subraveti, Z. Li, V. Prasad and A. Rajendran, Machine learning-based multiobjective optimization of pressure swing adsorption, *Ind. Eng. Chem. Res.*, 2019, **58**(44), 20412–20422.
- 75 A. Streb and M. Mazzotti, Performance limits of neural networks for optimizing an adsorption process for hydrogen purification and CO<sub>2</sub> capture, *Comput. Chem. Eng.*, 2022, **166**, 107974.
- 76 J. Beck, D. Friedrich, S. Brandani and E. S. Fraga, Multi-objective optimisation using surrogate models for the design of VPSA systems, *Comput. Chem. Eng.*, 2015, **82**, 318–329.
- 77 J. Du, H. Cao and Y. Li, *et al.*, Development of hybrid surrogate model structures for design and optimization of CO<sub>2</sub> capture processes: Part I. Vacuum pressure swing adsorption in a confined space, *Chem. Eng. Sci.*, 2024, **283**, 119379.
- 78 Y. Ureel, M. R. Dobbelaere and Y. Ouyang, *et al.*, Active machine learning for chemical engineers: a bright future lies ahead!, *Engineering*, 2023, **27**, 23–30.
- 79 K. Wang and A. W. Dowling, Bayesian optimization for chemical products and functional materials, *Curr. Opin. Chem. Eng.*, 2022, **36**, 100728.
- 80 J. A. Paulson and C. Tsay, Bayesian optimization as a flexible and efficient design framework for sustainable process systems, *Curr. Opin. Green Sustainable Chem.*, 2024, 100983.
- 81 L. D. González and V. M. Zavala, New paradigms for exploiting parallel experiments in Bayesian optimization, *Comput. Chem. Eng.*, 2023, **170**, 108110.
- 82 Z. Hao, A. Caspari, A. M. Schweidtmann, Y. Vaupel, A. A. Lapkin and A. Mhamdi, Efficient hybrid multiobjective optimization of pressure swing adsorption, *Chem. Eng. J.*, 2021, **423**, 130248.
- 83 A. Ward and R. Pini, Efficient bayesian optimization of industrial-scale pressure-vacuum swing adsorption processes for CO<sub>2</sub> capture, *Ind. Eng. Chem. Res.*, 2022, **61**(36), 13650–13668.
- 84 K. N. Pai, V. Prasad and A. Rajendran, Generalized, adsorbent-agnostic, artificial neural network framework for rapid simulation, optimization, and adsorbent screening of adsorption processes, *Ind. Eng. Chem. Res.*, 2020, **59**(38), 16730–16740.
- 85 H. Park, X. Yan and R. Zhu, *et al.*, A generative artificial intelligence framework based on a molecular diffusion model for the design of metal-organic frameworks for carbon capture, *Commun. Chem.*, 2024, **7**(1), 21.
- 86 I. Igashov, H. Stärk and C. Vignac, *et al.*, Equivariant 3D-conditional diffusion model for molecular linker design, *Nat. Mach. Intell.*, 2024, **6**(4), 417–427.
- 87 P. Ertl and A. Schuffenhauer, Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions, *Aust. J. Chem.*, 2009, **1**, 1–11.
- 88 C. W. Coley, L. Rogers, W. H. Green and K. F. Jensen, SCScore: synthetic complexity learned from a reaction corpus, *J. Chem. Inf. Model.*, 2018, **58**(2), 252–261.
- 89 H. Park, R. Zhu, E. Huerta, S. Chaudhuri, E. Tajkhorshid and D. Cooper, End-to-end AI framework for interpretable prediction of molecular and crystal properties, *Mach. Learn.: Sci. Technol.*, 2023, **4**(2), 025036.
- 90 F. Cipcigan, J. Booth, R. N. B. Ferreira, C. R. dos Santos and M. Steiner, Discovery of novel reticular materials for carbon dioxide capture using GFlowNets, *Digital Discovery*, 2024, **3**(3), 449–455.
- 91 S. Lee, B. Kim and H. Cho, *et al.*, Computational screening of trillions of metal-organic frameworks for high-performance methane storage, *ACS Appl. Mater. Interfaces*, 2021, **13**(20), 23647–23654.
- 92 C. Shen, M. Krenn, S. Eppel and A. Aspuru-Guzik, Deep molecular dreaming: Inverse machine learning for de-novo molecular design and interpretability with surjective representations, *Mach. Learn.: Sci. Technol.*, 2021, **2**(3), 03LT02.
- 93 C. Cleeton and L. Sarkisov, Inverse design of metal-organic frameworks using deep dreaming approaches, *Nat. Commun.*, 2025, **16**(1), 4806.
- 94 M. Krenn, F. Häse, A. Nigam, P. Friederich and A. Aspuru-Guzik, Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation, *Mach. Learn.: Sci. Technol.*, 2020, **1**(4), 045024.
- 95 B. Sanchez-Lengeling and A. Aspuru-Guzik, Inverse molecular design using machine learning: Generative models for matter engineering, *Science*, 2018, **361**(6400), 360–365.
- 96 Z. Deng and L. Sarkisov, Multi-scale computational design of metal-organic frameworks for carbon capture using machine learning and multi-objective optimization, *Chem. Mater.*, 2024, **36**(19), 9806–9821.
- 97 S. P. Collins, T. D. Daff, S. S. Piotrkowski and T. K. Woo, Materials design by evolutionary optimization of functional groups in metal-organic frameworks, *Sci. Adv.*, 2016, **2**(11), e1600954.



- 98 X. Zhang, K. Zhang and Y. Lee, Machine learning enabled tailor-made design of application-specific metal–organic frameworks, *ACS Appl. Mater. Interfaces*, 2019, **12**(1), 734–743.
- 99 X. Yang, J. Zhang, K. Yoshizoe, K. Terayama and K. Tsuda, ChemTS: an efficient python library for de novo molecular generation, *Sci. Technol. Adv. Mater.*, 2017, **18**(1), 972–976.
- 100 X. Zhang, K. Zhang, H. Yoo and Y. Lee, Machine learning-driven discovery of metal–organic frameworks for efficient CO<sub>2</sub> capture in humid condition, *ACS Sustainable Chem. Eng.*, 2021, **9**(7), 2872–2879.
- 101 H. Park, S. Majumdar, X. Zhang, J. Kim and B. Smit, Inverse design of metal–organic frameworks for direct air capture of CO<sub>2</sub> via deep reinforcement learning, *Digital Discovery*, 2024, **3**(4), 728–741.
- 102 Z. Yao, B. Sánchez-Lengeling and N. S. Bobbitt, *et al.*, Inverse design of nanoporous crystalline reticular materials with deep generative models, *Nat. Mach. Intell.*, 2021, **3**(1), 76–86.
- 103 B. Kim, S. Lee and J. Kim, Inverse design of porous materials using artificial neural networks, *Sci. Adv.*, 2020, **6**(1), eaax9324.
- 104 C. Cleeton, F. L. De Oliveira and R. F. Neumann, *et al.*, A process-level perspective of the impact of molecular force fields on the computational screening of MOFs for carbon capture, *Energy Environ. Sci.*, 2023, **16**(9), 3899–3918.
- 105 A. Rajendran, G. K. Shimizu and T. K. Woo, The challenge of water competition in physical adsorption of CO<sub>2</sub> by porous solids for carbon capture applications—a short perspective, *Adv. Mater.*, 2024, **36**(12), 2301730.
- 106 A. Raza, A. Sturluson, C. M. Simon and X. Fern, Message passing neural networks for partial charge assignment to metal–organic frameworks, *J. Phys. Chem. C*, 2020, **124**(35), 19070–19082.
- 107 S. Kancharlapalli, A. Gopalan, M. Haranczyk and R. Q. Snurr, Fast and accurate machine learning strategy for calculating partial atomic charges in metal–organic frameworks, *J. Chem. Theory Comput.*, 2021, **17**(5), 3052–3064.
- 108 G. Zhao and Y. G. Chung, PACMAN: A robust partial atomic charge predictor for nanoporous materials based on crystal graph convolution networks, *J. Chem. Theory Comput.*, 2024, **20**(12), 5368–5380.
- 109 J. Luo, O. B. Said and P. Xie, *et al.*, MEPO-ML: A robust graph attention network model for rapid generation of partial atomic charges in metal-organic frameworks, *npj Comput. Mater.*, 2024, **10**(1), 224.
- 110 R. Batra, C. Chen, T. G. Evans, K. S. Walton and R. Ramprasad, Prediction of water stability of metal–organic frameworks using machine learning, *Nat. Mach. Intell.*, 2020, **2**(11), 704–710.
- 111 A. Nandy, C. Duan and H. J. Kulik, Using machine learning and data mining to leverage community knowledge for the engineering of stable metal–organic frameworks, *J. Am. Chem. Soc.*, 2021, **143**(42), 17535–17547.
- 112 A. Nandy, G. Terrones, N. Arunachalam, C. Duan, D. W. Kastner and H. J. Kulik, MOFSimplify, machine learning models with extracted stability data of three thousand metal–organic frameworks, *Sci. Data*, 2022, **9**(1), 74.
- 113 G. G. Terrones, S.-P. Huang, M. P. Rivera, S. Yue, A. Hernandez and H. J. Kulik, Metal–organic framework stability in water and harsh environments from data-driven models trained on the diverse WS24 data set, *J. Am. Chem. Soc.*, 2024, **146**(29), 20333–20348.
- 114 S. Zhang, Z. Wang, H. Gao and T. Zhou, A multitask learning model for predicting various types of metal–organic framework stability, *Ind. Eng. Chem. Res.*, 2025, **64**(29), 14576–14589.
- 115 P. Z. Moghadam, Y. G. Chung and R. Q. Snurr, Progress toward the computational discovery of new metal–organic framework adsorbents for energy applications, *Nat. Energy*, 2024, **9**(2), 121–133.
- 116 M. Voršilák, M. Kolář, I. Čmelo and D. Svozil, SYBA: Bayesian estimation of synthetic accessibility of organic compounds, *Aust. J. Chem.*, 2020, **12**(1), 35.
- 117 A. Thakkar, V. Chadimová, E. J. Bjerrum, O. Engkvist and J.-L. Reymond, Retrosynthetic accessibility score (RAScore)—rapid machine learned synthesizability classification from AI driven retrosynthetic planning, *Chem. Sci.*, 2021, **12**(9), 3339–3349.
- 118 C. G. Livas, P. N. Trikalitis and G. E. Froudakis, MOFSynth: A computational tool toward synthetic likelihood predictions of MOFs, *J. Chem. Inf. Model.*, 2024, **64**(21), 8193–8200.
- 119 A. M. Wright, M. T. Kapelewski, S. Marx, O. K. Farha and W. Morris, Transitioning metal–organic frameworks from the laboratory to market through applied research, *Nat. Mater.*, 2025, **24**(2), 178–187.
- 120 D. W. Siderius, V. K. Shen, R. D. Johnson III and R. D. van Zee, *NIST/ARPA-E Database of Novel and Emerging Adsorbent Materials*, National Institute of Standards and Technology.
- 121 D. Ongari, L. Talirz, K. M. Jablonka, D. W. Siderius and B. Smit, Data-driven matching of experimental crystal structures and gas adsorption isotherms of Metal–Organic frameworks, *J. Chem. Eng. Data*, 2022, **67**(7), 1743–1756.
- 122 P. Iacomi and P. L. Llewellyn, Data mining for binary separation materials in published adsorption isotherms, *Chem. Mater.*, 2020, **32**(3), 982–991.
- 123 S. Park, B. Kim, S. Choi, P. G. Boyd, B. Smit and J. Kim, Text mining metal–organic framework papers, *J. Chem. Inf. Model.*, 2018, **58**(2), 244–251.
- 124 L. T. Glasby, K. Gubsch and R. Bence, *et al.*, DigiMOF: a database of metal–organic framework synthesis information generated via text mining, *Chem. Mater.*, 2023, **35**(11), 4510–4524.
- 125 A. J. White, M. Gibaldi, J. Burner, R. A. Mayo and T. K. Woo, High structural error rates in “computation-ready” MOF databases discovered by checking metal oxidation states, *J. Am. Chem. Soc.*, 2025, **147**(21), 17579–17583.
- 126 X. Jin, K. M. Jablonka, E. Moubarak, Y. Li and B. Smit, MOFChecker: A package for validating and correcting



- metal–organic framework (MOF) structures, *Digital Discovery*, 2025, **4**, 1560–1569.
- 127 A. Deshwal, C. M. Simon and J. R. Doppa, Bayesian optimization of nanoporous materials, *Mol. Syst. Des. Eng.*, 2021, **6**(12), 1066–1086.
- 128 N. Gantzer, A. Deshwal, J. R. Doppa and C. M. Simon, Multi-fidelity Bayesian optimization of covalent organic frameworks for xenon/krypton separations, *Digital Discovery*, 2023, **2**(6), 1937–1956.
- 129 X. Zhang, T. Zhou and K. Sundmacher, Integrated metal–organic framework and pressure/vacuum swing adsorption process design: Descriptor optimization, *AIChE J.*, 2022, **68**(2), e17524.
- 130 X. Zhang, T. Zhou and K. Sundmacher, Integrated metal–organic framework (MOF) and pressure/vacuum swing adsorption process design: MOF matching, *AIChE J.*, 2022, **68**(9), e17788.
- 131 C. Charalambous, E. Moubarak and J. Schilling, *et al.*, A holistic platform for accelerating sorbent-based carbon capture, *Nature*, 2024, **632**(8023), 89–94.
- 132 Z. Wang, T. Zhou and K. Sundmacher, Data-driven integrated design of solvents and extractive distillation processes, *AIChE J.*, 2023, **69**(12), e18236.
- 133 P. Daoutidis, J. H. Lee and S. Rangarajan, *et al.*, Machine learning in process systems engineering: Challenges and opportunities, *Comput. Chem. Eng.*, 2024, **181**, 108523.

