


Article

Combining Computational Screening and Machine Learning to Predict Metal–Organic Framework Adsorbents and Membranes for Removing CH₄ or H₂ from Air

Huilin Li ¹, Cuimiao Wang ¹, Yue Zeng ¹, Dong Li ¹, Yaling Yan ^{1,*}, Xin Zhu ^{1,*} and Zhiwei Qiao ^{1,2,*} 

¹ Guangzhou Key Laboratory for New Energy and Green Catalysis, School of Chemistry and Chemical Engineering, Guangzhou University, Guangzhou 510006, China

² Joint Institute of Guangzhou University & Institute of Corrosion Science and Technology, Guangzhou University, Guangzhou 510006, China

* Correspondence: yaling@e.gzhu.edu.cn (Y.Y.); zhux@gzhu.edu.cn (X.Z.); zqiao@gzhu.edu.cn (Z.Q.)

Abstract: Separating and capturing small amounts of CH₄ or H₂ from a mixture of gases, such as coal mine spent air, at a large scale remains a great challenge. We used large-scale computational screening and machine learning (ML) to simulate and explore the adsorption, diffusion, and permeation properties of 6013 computation-ready experimental metal–organic framework (MOF) adsorbents and MOF membranes (MOFMs) for capturing clean energy gases (CH₄ and H₂) in air. First, we modeled the relationships between the adsorption and the MOF membrane performance indicators and their characteristic descriptors. Among three ML algorithms, the random forest was found to have the best prediction efficiency for two systems (CH₄/(O₂ + N₂) and H₂/(O₂ + N₂)). Then, the algorithm was further applied to quantitatively analyze the relative importance values of seven MOF descriptors for five performance metrics of the two systems. Furthermore, the 20 best MOFs were also selected. Finally, the commonalities between the high-performance MOFs were analyzed, leading to three types of material design principles: tuned topology, alternative metal nodes, and organic linkers. As a result, this study provides microscopic insights into the capture of trace amounts of CH₄ or H₂ from air for applications involving coal mine spent air and hydrogen leakage.

Keywords: metal–organic framework; membrane; molecular simulation; computational screening; machine learning



Citation: Li, H.; Wang, C.; Zeng, Y.; Li, D.; Yan, Y.; Zhu, X.; Qiao, Z. Combining Computational Screening and Machine Learning to Predict Metal–Organic Framework Adsorbents and Membranes for Removing CH₄ or H₂ from Air. *Membranes* **2022**, *12*, 830. <https://doi.org/10.3390/membranes12090830>

Academic Editor: Anthony G. Dixon

Received: 23 July 2022

Accepted: 23 August 2022

Published: 25 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the development and advancement of society and the economy, the demand for fossil energy fuels is increasing with each passing day, leading to a widespread problem of CH₄ leakage in fossil energy extraction, production, transportation, and application. For instance, leaking in underground boreholes for gas extraction can diminish CH₄ production and contaminate groundwater, and CH₄ also volatilizes in gas stations. More noteworthy is that the total amount of coal mine methane (coal mine gas with methane concentration below 0.75%) is huge. During the mining process, the total amount of methane emitted is about 28 billion m³/a, which is equivalent to 420 million tons of CO₂, accounting for 8% of the total anthropogenic methane emissions. Moreover, combined with its low viscosity and the technical difficulty of its use, this methane has had to be exhausted entirely into the atmosphere for long periods, causing significant greenhouse gas pollution. Apart from fossil fuels, natural gas, hydrogen energy, and hydrogen fuel cells are also developing rapidly owing to their high green energy value. However, one of the main factors hindering the commercialization of hydrogen-fueled vehicles is the safety and cost caused by the leakage of H₂. Meanwhile, CH₄ and H₂ are clean energy carriers with high calorific value and low carbon emissions. It would be progress if we could capture this part of the H₂ and CH₄ free in the air. Therefore, determining how to separate and capture the relatively small

amounts of CH₄ and H₂ from air (O₂ and N₂) has become the focus of research. Exploiting diverse new adsorbents or separation membranes is the key for the safe and cost-effective storage and separation of gases (CH₄ and H₂).

In recent years, a newly developed category of porous coordination polymer materials has attracted significant research attention and received widespread attention in scientific research and industry. Obtained by self-assembly of metal ions and organic ligands, metal–organic frameworks (MOFs) [1–3] not only have crystalline structures similar to the regular pores of zeolite molecular sieves but also have higher specific surface areas than conventional porous materials, and organic components make them both designable and tailorable, with adjustable pore sizes and easy functionalization of the channel surfaces. Based on these structural features, which act as adsorbents, MOF materials can achieve high-density energy storage of clean fuel gases. MOFs can also generate differential interactions for different gas molecules to achieve economical and energy-efficient separation of gases, and they are therefore widely used in gas adsorption and separation [4–7], storage [8], catalysis [9,10], drug transport [11], and sensing applications [12], especially in the field of gas adsorption and separation in which a number of breakthroughs have been accomplished. Chang [13] showed in breakthrough experiments that a microporous MOF (SBMOF-1) allowed good separation of methane from N₂. Xu [14] found it effective to combine B-substitution and Li-decoration or C₄₈B₁₂Li insertion as a strategy to improve the capacity of CO₂ to absorb H₂ and methane and to separate it from CO₂/CH₄ and CO₂/H₂ mixtures. This multiple modification strategy opened a door for the production of porous nanomaterials with higher gas adsorption and separation capabilities. Kang [15] found in his study that MOFJUC-150 membranes have a significant preferential permeability to H₂ compared to other gas molecules. At ambient temperature, the selectivity factors of the membranes for H₂/CH₄, H₂/N₂, and H₂/CO₂ were 26.3, 17.1, and 38.7. MOFs have also been reported to possess ultra-high surface areas, especially for a series of CH₄/H₂ separations, with higher selective adsorption capacities than conventional porous materials, indicating that such materials are the most attractive adsorbent materials for CH₄/H₂ separation [16,17]. Eddaoudi and co-workers [18] developed several novel MOF materials for efficient gas storage, and one of them, a hybrid material, could carry out the cost-effective storage of methane.

MOFs have experienced rapid development over the past decade. In addition to being used as adsorbents, the properties of MOF membrane materials make them suitable for gas separation. Kang et al. [19] summarized the advances in improving the performance of MOF membranes (MOFMs) in recent years, including practical application problems faced by the design and growth of MOFs. Separation processes by using membrane materials have shown to be efficient, have low energy costs, can be easily performed [20], and have advantages in terms of environmental safety and scalability. Conventional materials, such as carbon, zeolites, and polymers, have been explored as membranes for the separation of natural gas. However, the number of zeolite structures and the diversity of zeolite membranes are restricted, and it is difficult to control the precise pore size and pore function of zeolites [21]. Polymeric membranes, however, are often subject to a balance between the permeability and selectivity (called Robson's upper limit); that is, the increase in the permeability in conventional membrane materials occurs at the cost of selectivity [22]. Nevertheless, MOF membrane materials show strong potential. Hou et al. [16] reported a 10-fold higher separation coefficient, up to 25, for ZIF-722-8 membranes in CO₂/methane mixtures. Liu's group [23] found that fcu-MOF membranes exhibited good permeability and selectivity for the separation of H₂/CO₂/N₂, CO₂/CH₄, and N₂/CH₄ mixtures. Some MOF membrane materials even exceeded Robson's upper limit. For example, Wang [24] developed the ZIF-62MOF glass membrane with separation coefficients of 50.7, 34.5, and 36.6 for H₂/CH₄, CO₂/N₂, and CO₂/CH₄ mixtures, respectively.

To date, although a large number of MOFs have been synthesized and reported, the number is almost infinite due to the large number of possible metal ions and organic linkers. Consequently, it is inefficient to filter MOFs from a large database for a specific

application for which the high-throughput computational screening (HTCS) method based on molecular simulations provides a suitable alternative. Numerous studies [25–28] have shown that the method of discovering high-performance target materials and mining the quantitative structure–property relationships (QSPRs) from the large number of MOFs provide is effective for selecting superior MOFs on a large scale. However, many inefficient computations are performed during high-throughput computing, resulting in a considerable waste of computational resources and valuable research time. Recently, the emergence of machine learning (ML) has made up for this shortcoming. ML has been gradually applied in many fields, such as material discovery, structure analysis, property prediction, and reverse design, and has shown high potential in materials research [6,29–32]. Li et al. [33] summarized the latest advances in the use of MOFs for gas storage from the three aspects of H₂, CH₄, and C₂H₂. In addition to gas storage, it highlights some of the significant advances made by MOF materials in the separation of important gases in recent years. According to Shi et al.'s work, ML models were trained to be two to three orders of magnitude faster than HTCS. The combination of ML and simulation techniques is an effective way to find the optimal MOF quickly and with the greatest probability [6]. For example, Yan et al. [34] screened the dynamic adsorption of O₂ and N₂ in 6013 computation-ready experimental MOFs (CoRE-MOFs) by using ML and large-scale calculations. They also systematically analyzed the influences of all the metals in the periodic table on the performances of the materials and proposed six rational design criteria for adsorbent materials, successfully designing a series of materials with superior performances. Recently, Jiang and co-workers [35] deployed a hierarchical approach using molecular simulations and machine learning to rapidly screen 100,000+ MOFs that were synthesized experimentally for C₃ separation, and they used trained ML models for rapid screening of other MOF datasets (e.g., experimental Cambridge Structural Database (CSD) MOFs and hypothetical MOFs). For the CSD MOFs, the out-of-sample predictions closely coincided with the simulation results, which indicated good transferability of the ML model from the CoRE-MOFs to CSD MOFs. Furthermore, nine CSD MOFs that exhibited better separation performances than the best performing CoRE-MOFs were revealed. Rosen et al. [36] trained an ML model on a quantum database (QMOF) with 14,000 MOFs to discover MOFs with structural properties of the target electrons in a rapid manner. Through reconstruction of the typical track of failed experiments, Moosavi et al. [37] reported an ML approach to obtaining chemical intuition and successfully searched for the optimal synthesis conditions, which yielded the highest surface area of HKUST-1 reported to date. He also illustrated the importance of quantifying this intuition for the synthesis of new materials. Azar [38] predicted the H₂ permeability and H₂/N₂ selectivity from 3765 different types of MOFMs, and the results demonstrated that MOFMs had a high H₂ permeability, 2.5×10^3 to 1.7×10^6 barrer. Qiao et al. [26] identified five optimized MOFs at 298 K and 10 bar by simulating the adsorption, diffusion, and permeation of CO₂/N₂/CH₄ mixtures in 24 samples using a computational study of the high-throughput screening of 137,953 MOFs. Qiao also selected 4764 CoRE-MOFs for membrane separation of ternary gas mixtures (CO₂/N₂/CH₄) at 298 K and 10 bar and finally identified seven PLDs of the best MOFs with 2.91–3.26 Å and pore size distribution (PSD%_(2.4 to 3.5 Å)) of 48.2–64.1% [39]. Recently, Bai et al. [40] targeted computational screening 6013 MOFMs for H₂ separation and explored the relationship between material characteristics and properties from the perspective of ML. The main difference with Bai's research is the separation of trace amounts of CH₄ (or H₂) and ternary systems in this work. All these again confirmed that the synergistic use of ML and HTCS is an effective way to achieve faster, better predictions and probability maximization of high-performance MOFs or MOFMs.

Gas separation using MOFs is generally divided into two categories: equilibrium-based gas separations and kinetic-based gas separations [41]. In equilibrium-based gas separations where MOFs are used as adsorbents, its selectivity of equilibrium-based separations was controlled by the affinity of MOFs for adsorption of one gas relative to another. In kinetic-based separations, the selectivity is controlled by the combination of adsorp-

tion and diffusion, being determined by the different transport rates of the gas species through the membrane pores, and MOFs are used as membranes [42]. Admittedly, the use of membrane separation for processing trace gases requires a larger driving force and increases cost; however, MOFMs have certain properties and their applicability to specific sites that are superior to MOF adsorbents and other types of membrane materials. Firstly, the present simulations were performed under atmospheric pressure and temperature conditions, which are applicable to general situations. Secondly, MOF adsorbents are less stable in aqueous or high heat environments, while membranes have stronger stability, e.g., ZIF-90 membranes show good steadiness in the presence of steam [43]. Therefore, membrane materials may be more advantageous when operating in liquid or high temperature environments. Gokay Avci et al. [44] also proposed that the separation of H₂ is more advantageous in membrane applications than CO₂ because H₂ has a higher molar fraction and a smaller kinetic diameter, which allows H₂ to penetrate smaller membrane pores more quickly. In addition, MOFMs are better equipped for further modification and refinement of the material, e.g., as polymer-filled particles in polymers to improve polymer separation performance, when making MOF-based MMM is much more economical than making pure MOF membranes, since the former requires only a small amount of MOF [43]. MOFs have good potential in pressure-driven membrane processes, as well as in selective filtration or permeation-driven membrane processes. After comprehensive consideration, in some specific occasions, MOF membranes will have better performance and more economical applications. In summary, it is also necessary to develop materials with good membrane separation performance.

The purpose of this work was to investigate the adsorption, diffusion, and permeation properties of 6013 MOF adsorbents and MOFMs for the capture of clean energy gases (CH₄ and H₂) in air using large-scale computational screening and ML. In Section 2, we described atomic models of CoRE-MOFs/MOFMs and gases, simulation methodologies, and machine learning principles. In Section 3, we discussed the relationships between the adsorbent performance metrics and their characteristic descriptors. With the help of ML techniques, we found the best ML model for a particular performance metric aimed at CH₄/O₂ + N₂ and H₂/O₂ + N₂ mixtures. Furthermore, this algorithm was applied to quantify the relative importance of the performance metrics for each MOF descriptor. Finally, high-performance MOF materials for different systems were identified, and commonalities between the high-performance MOFs were evaluated to propose design principles.

2. Models and Methods

2.1. Model

In this work, we used molecular simulation to screen 6013 computation-ready, experimental MOFs (CORE-MOFs) for their ability to capture energy gases (CH₄ and H₂) from air. All crystal structures were taken from the Cambridge Crystallographic Data Centre (CCDC), whose parameters were compiled and validated by Chung et al. [45]. Each MOF is described by five structural descriptors (largest cavity diameter (LCD, Å), void fraction (ϕ), volumetric surface area (VSA, m²/cm³), pore limiting diameter (PLD, Å), density (ρ , kg/m³)) and two energy descriptors (heat of adsorption (Q_{st} , kJ/mol) and Henry's constant (K , mol/kg/Pa)). Specifically, N₂ with a diameter of 3.64 (Å) and He with a diameter of 2.58 (Å) were used as probes in RASPA to calculate VSA and ϕ . Then, LCD and PLD were calculated using the Zeo++ [46] package. K was simulated in RASPA using the NTV-MC Scheme (N is the number of molecules, T is the temperature, V is the volume, and MC denotes Monte Carlo). The structural atoms of MOFs are described by the Lennard–Jones (LJ) potential and electrostatic potential, and all the LJ potential parameters are derived from the universal force field (UFF) [47], which is listed in Table S1. The atomic charges of MOFs can be quickly calculated using the MOF electrostatic-potential-optimized charge equilibration (MEPO-Qeq) method, which helps to accurately screen MOFs to adsorb specific gases. The force field parameters of CH₄, H₂, O₂, and N₂ are derived from the TraPPE force field, which is listed in Table S2. A number of studies have verified that the UFF

force field, the MEPO-Qeq charge algorithm, and the TraPPE [48] force field can accurately predict the adsorption and diffusion of gases in different MOFs materials [49,50].

2.2. Simulation Method

We simulated the adsorption and diffusion characteristics of 6013 CoRE-MOFs for airborne capture of energy gases (CH_4 and H_2) using grand canonical MC (GCMC) and Molecular Dynamics (MD) at 1 bar and 298 K. The interactions between MOFs and adsorbent molecules were calculated by the Lorentz–Berthelot rule, and periodic boundaries were imposed in the three-dimensional system to simulate the cell expansion along the three-dimensional direction to at least 24 Å. To calculate the LJ interactions, the spherical truncation radius for the long-range correction was set to 12 Å. The electrostatic interactions between the frame and gas molecules and between the gas molecules were calculated using the summation of Ewald [51]. In each MOF, the MC simulation was run for 100,000 cycles. The first 50,000 cycles were used to equilibrate the simulated system, and the second 50,000 cycles were used to calculate the average value of the pendulum amplitude. Each cycle consisted of tests for experiment with n (n : number of adsorbed molecules including translation, rotation, regeneration, and exchange). One GCMC [52] and MD simulation were run independently for each of the 6013 MOFs. In the MD simulations, for each ternary gas mixture system in each MOF the MD duration was 7 ns, and the last 5 ns was used for production. It was found that further increases in the number of cycles and MD duration had little effect on the simulation results. All simulations were run under the RASPA package.

2.3. Machine Learning

The rapidly evolving field of ML-assisted materials research and development has made it possible to combine materials databases and machine learning methods to drive materials discovery and design and predict material properties. The machine learning part of this work was performed using a Python-based automated machine learning development tool (tree-based pipeline optimization tool (TPOT)), a fast model selection and tuning method based on genetic algorithms. In addition, two machine learning algorithms (decision tree (DT) and random forest (RF)) from the Scikit-learn package [53] in Python 3.9 were used for each performance metric (adsorption selectivity (S_{ads}), diffusion selectivity (S_{diff}), diffusion coefficient (D), permeability (P), and permselectivity (S_{perm})) for each sorbent (CH_4 and H_2). It is worth mentioning that the first three performance indicators correspond to MOF adsorbent applications, and the last two performance indicators correspond to MOF membrane applications. More details of the algorithm are shown in Figures S6–S8. For each ML method, the data set was randomly divided into a training set and a test set, where 80% was used for training the model and the rest for testing. We used k -fold cross-validation ($k = 5$) for model construction to reduce the effect of data partitioning during learning, and we set a fixed random seed to ensure the reproducibility of the results. Each algorithm was repeated five times in our work, and the final results were averaged for each performance metric prediction. In addition, the accuracy of the models was evaluated using the Pearson correlation coefficient (R), mean absolute error (MAE), and root mean square error (RMSE). More details can be found in the Supporting Materials (SM).

3. Results and Discussion

In this work, 6013 CoRE-MOFs were used to adsorb and separate mixtures of energy gases (CH_4 or H_2 at 1000 ppm) with air ($\text{N}_2:\text{O}_2 = 78:21$), an initial concentration that mimics the CH_4 concentration of spent coal mine air and the concentration of a small amount of H_2 leaking inside a confined hydrogen fuel cell. This separation aim is the capture of CH_4 or H_2 . With this micro level at atmospheric pressure (1 bar), it can also be theoretically shown that it makes sense to the separation of CH_4 or H_2 in air. In order to design superior adsorbents or membranes, first, the relationships between seven structural/energetic descriptors (LCD, ϕ , VSA, PLD, ρ , Q_{st} , and K) of 6013 CoRE-MOFs and the static adsorption, diffusion performance, and permeation properties (D , S_{ads} , S_{diff} , P , and S_{perm}) were explored by

univariate analysis for two systems ($\text{CH}_4/\text{N}_2 + \text{O}_2$ and $\text{H}_2/\text{N}_2 + \text{O}_2$). The relationships were obtained, and the preliminary influence characteristics affecting the adsorption and diffusion of CH_4 (or H_2) in air were obtained. Then, the complex structure–property relationships were further explored by machine learning methods (TPOT, DT, and RF). Finally, the best candidates were identified for different applications, and the corresponding design principles of adsorbent/membrane materials were proposed.

3.1. Adsorption and Diffusion

Understanding the correlation between the diffusion coefficients of CH_4 (or H_2) for N_2 and O_2 , adsorption selectivity, and diffusion selectivity as well as the structure/energy descriptors of the MOFs in terms of both adsorption and diffusion separation properties can help to uncover materials with potential for specific applications. Figures S1 and S2 show highly similar trends in the structure–property relationships in both systems, so only the $\text{CH}_4/\text{N}_2 + \text{O}_2$ hybrid system was elaborated. Figures S1a–d,g clearly show that D , S_{ads} , and S_{diff} first increased and finally leveled off as the five descriptors (LCD, ϕ , VSA, PLD, and K) increased. When the pore size was small, the pore walls of different framework materials exhibited strong and weak adsorption of gases, and high variability in the selectivity occurred. As the pore size increased, both N_2 and O_2 entered the MOFs in large amounts, and the selectivity decreased dramatically and eventually converged to one for both the $\text{CH}_4/\text{N}_2 + \text{O}_2$ and $\text{H}_2/\text{N}_2 + \text{O}_2$ systems. The ϕ , VSA, and PLD showed similar relationships, which was consistent with the previously reported trends [25,54]. It is worth noting that in Figure S1e, three performance metrics, S_{ads} , S_{diff} , and D , all tended to decrease as ρ increased. This was because a higher density of an MOF corresponded to a very dense internal space. As shown in Figure S1f, there were no significant trends of the three performance indicators, S_{ads} , S_{diff} , and D , with Q_{st} , which indicated the insignificant role of Q_{st} .

In addition to adsorption and diffusion selectivity, Henry's constant of CH_4 (or H_2) reflects the adsorption performance of CH_4 (or H_2) in an infinitely dilute state and helps to explain the trapping performances of the MOFs in air with very low CH_4 (or H_2) concentrations. Figure 1 clearly shows the trend of the K values with PLD for the four gases. The trends of K for N_2 and O_2 were almost the same. In general, K decreased with $\text{CH}_4 > \text{N}_2 (\approx \text{O}_2) > \text{H}_2$ because CH_4 had the highest affinity for most MOFs, while H_2 had the weakest affinity. This is the reason that CH_4 and H_2 could be separated in $\text{CH}_4/\text{O}_2/\text{N}_2$ and $\text{H}_2/\text{O}_2/\text{N}_2$ mixtures, respectively. Then, the quantitative relationship $\log D = a\text{PLD} - b$ was established based on the diffusion coefficients of CH_4 , H_2 , N_2 , and O_2 at an infinite dilution and with PLD, as shown in Figure 2a,b. For most of the MOFs, $D_{\text{H}_2} > D_{\text{O}_2} > D_{\text{N}_2} > D_{\text{CH}_4}$, and according to Table S3, this was because the gas molecules with smaller kinetic diameter diffused more quickly than their larger counterparts. This conforms to the conclusions reached in previous reports [26,34]. The above univariate analysis could only tentatively determine the relationship between the individual parameters and performance. We further utilized the ML algorithms to systematically and comprehensively analyze the integrated structure–performance relationship for the MOFs.

3.2. Permeation

To explore the separation performances of MOF membranes, two commonly used membrane separation performance metrics were further calculated: the permeability P ($P = K \times D$) and the permeation selectivity S_{perm} ($S_{\text{perm}} = S_{\text{ads}} \times S_{\text{diff}}$). An evaluation method [26,40,55] has been shown to be applicable for calculating MOF membrane performances in multiple studies. Figure 2c,d show the P versus PLD plots of CH_4 , H_2 , N_2 , and O_2 for 6013 MOFs. As shown in Figure 2a,b, the variation of P with PLD was essentially similar to that of D with PLD because $P = KD$, and K was almost independent of PLD. Therefore, there was also a quantitative relationship between $\log P$ and PLD: $\log P = c\text{PLD} - d$. Slopes c for CH_4 , N_2 , O_2 , and H_2 were 1.484, 1.134, 0.827, and 0.631, respectively. It is clear that the magnitude of the slope was positively correlated with the molecular dynamic diameters of

the four gases. The permeability of CH_4 was most affected by PLD, and the permeability of H_2 was least affected by PLD because CH_4 had the largest molecular diameter of the four gases, while H_2 had the smallest. The effects of other parameters, LCD, ϕ , VSA, and PLD, on P and S_{perm} are shown in Figures S3 and S4. For the $\text{CH}_4/\text{O}_2 + \text{N}_2$ system, the trends of the structure–property relationships were almost the same as those in Figures S1 and S2. In contrast, for $\text{H}_2/\text{O}_2 + \text{N}_2$, Figure S4 shows similar variations of P with the seven structural parameters to those of the CH_4 system. However, S_{perm} showed a completely opposite trend with the structural parameters LCD, ϕ , VSA, PLD, and K . As shown in Table S3, this was mainly because O_2 and N_2 had higher molecular dynamics diameters than H_2 and smaller kinetic diameter than CH_4 . Small-pore MOFMs had different permeabilities to different size gas molecules, and this difference in permeability eventually led to a trend opposite to that of the permeation selectivity. S_{perm} decreased with an increase in any of the above five MOF descriptors. When PLD was small, the H_2 molecules could pass through the MOFMs, whereas the N_2 and O_2 molecules could hardly pass through. Thus, it possessed a high S_{perm} . When PLD increased slowly, the other two gas molecules could pass through the MOFMs freely; thus, S_{perm} decreased significantly and eventually tended to one. S_{perm} had a weakly decreasing trend with increasing ρ , as shown in Figure S4e. This was because high-density MOFMs had only small free spaces that were permeable to gas molecules, which led to low permeability values. However, some of the discrete points exhibited high P_{H_2} values at $>2000 \text{ kg/m}^3$. These MOFs had relatively large free spaces, and they were composed of very heavy metal atoms (e.g., gold, platinum, and uranium), consistent with the findings of previous studies [39,40]. A variation trend of S_{perm} with Q_{H_2} is not evident in Figure S4f, and most of the MOFMs exhibited moderate Q_{H_2} values (8–12 kJ/mol). The greater the gas adsorption heat of the material was, the stronger the forces between the H_2 gas molecules and the surface of the adsorbent material were, and the better the selectivity was. However, at the same time, energy cost required for desorption was higher, so a good adsorbent material should have a suitable adsorption heat. From the above analysis, it can be seen that it is not possible to obtain MOFMs with excellent performances based on a single descriptor. To understand the complex relationships between multiple descriptors and each performance metric, ML was further used to analyze the constitutive relationships of the MOFMs.

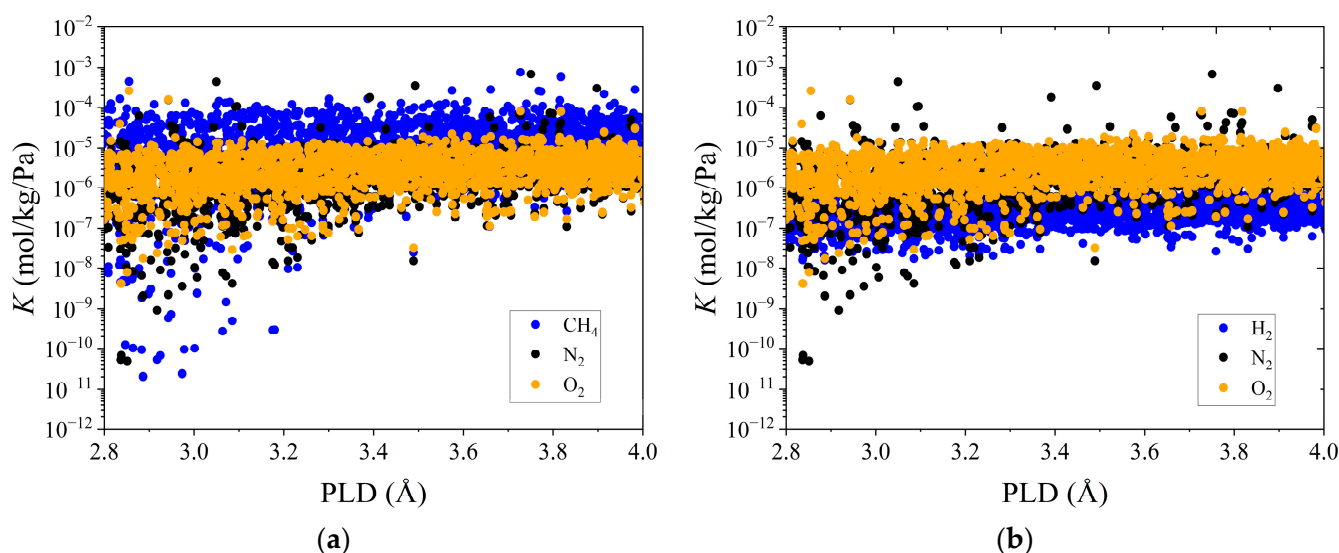


Figure 1. (a) Henry's constant K vs. PLD for CH_4 , N_2 , and O_2 in 6013 CoRE-MOFs; (b) Henry's constant K vs. PLD for H_2 , N_2 , and O_2 in 6013 CoRE-MOFs.

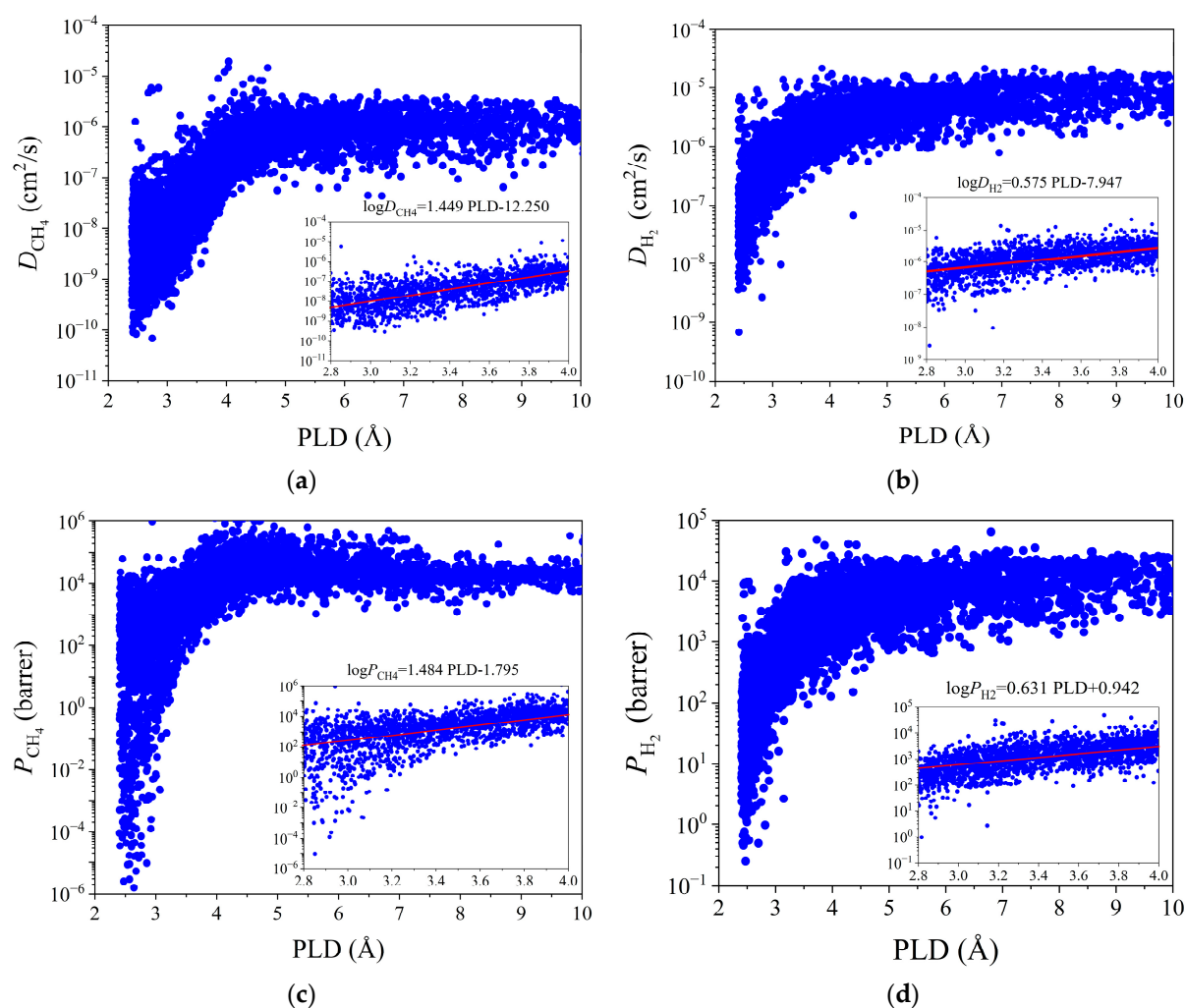


Figure 2. Diffusion coefficient D and permeability P vs. PLD for CH_4 and H_2 in 6013 CoRE-MOFs: (a) D_{CH_4} -PLD; (b) D_{H_2} -PLD; (c) P_{CH_4} -PLD; (d) P_{H_2} -PLD.

3.3. Machine Learning

Through univariate analysis, the relationships between the MOF structure/energy descriptors and adsorption/membrane separation performance were preliminarily obtained. In order to further understand the deeper relationships between the material performance and structure, especially the comprehensive effects of various MOFs descriptors on the performance and the ranking of the influence effects, we introduce an automatic machine learning development tool, TPOT. TPOT uses the structural characteristics and target performances of the MOFs as inputs and automatically generates models to achieve regression predictions of the target performances. At the same time, the best code pipeline for different target performances can be derived for further optimization and learning. In addition, two commonly used ML algorithms were employed: DT and RF. DT is a very commonly used ML algorithm. It achieves regression predictions by building a binary tree. The decision tree model is easy to implement and has strong interpretability. The RF algorithm is composed of multiple decision trees, which improves the generalization and fault tolerance of DT.

In order to find a machine learning method suitable for this system and obtain better prediction results, we used the TPOT, DT, and RF algorithms to predict each performance index (D , S_{ads} , S_{diff} , P , and S_{perm}). All the results are shown in Table 1 and Figure 3. It can be seen from Table 1 that the three algorithms yielded good predictions ($R \geq 0.85$) for all the performances of the two systems, except that the prediction effect of H_2 on the diffusion selectivity of N_2 and O_2 was not good. The TPOT and RF models had good prediction

abilities for the prediction of the adsorption selectivity S_{ads} and permeability P , and their R values were above 0.97 (see Figure 3). It is worth noting that, to narrow down a poor fitting effect caused by data span, we used logarithms to analyze $S_{\text{ads(H2/O2+N2)}}$, $S_{\text{diff(H2/O2+N2)}}$, $P_{\text{(CH4/O2+N2)}}$, and $S_{\text{perm(CH4/O2+N2)}}$. It can be seen from Figure 3 that the predictions of the RF model for the CH_4 adsorption selectivity and gas permeability were the closest to the simulation values and had the best prediction effect. The R values of $S_{\text{ads(CH4/O2+N2)}}$ and $P_{\text{(CH4/O2+N2)}}$ on the test set reached 0.96 and 0.99, the MAE values were 0.36 and 0.26, and the RMSE values were 0.64 and 0.36, respectively. In terms of adsorption and separation, the prediction effects of the three ML models on the performance indicators of the two systems were in the order of $S_{\text{ads}} > S_{\text{diff}} > D$. However, in terms of membrane performance, for the three ML models, although the R values of S_{perm} and P in the test set were similar, the MAE and RMSE of P were more than 1000 times those of S_{perm} , which was caused by the characteristics of the data.

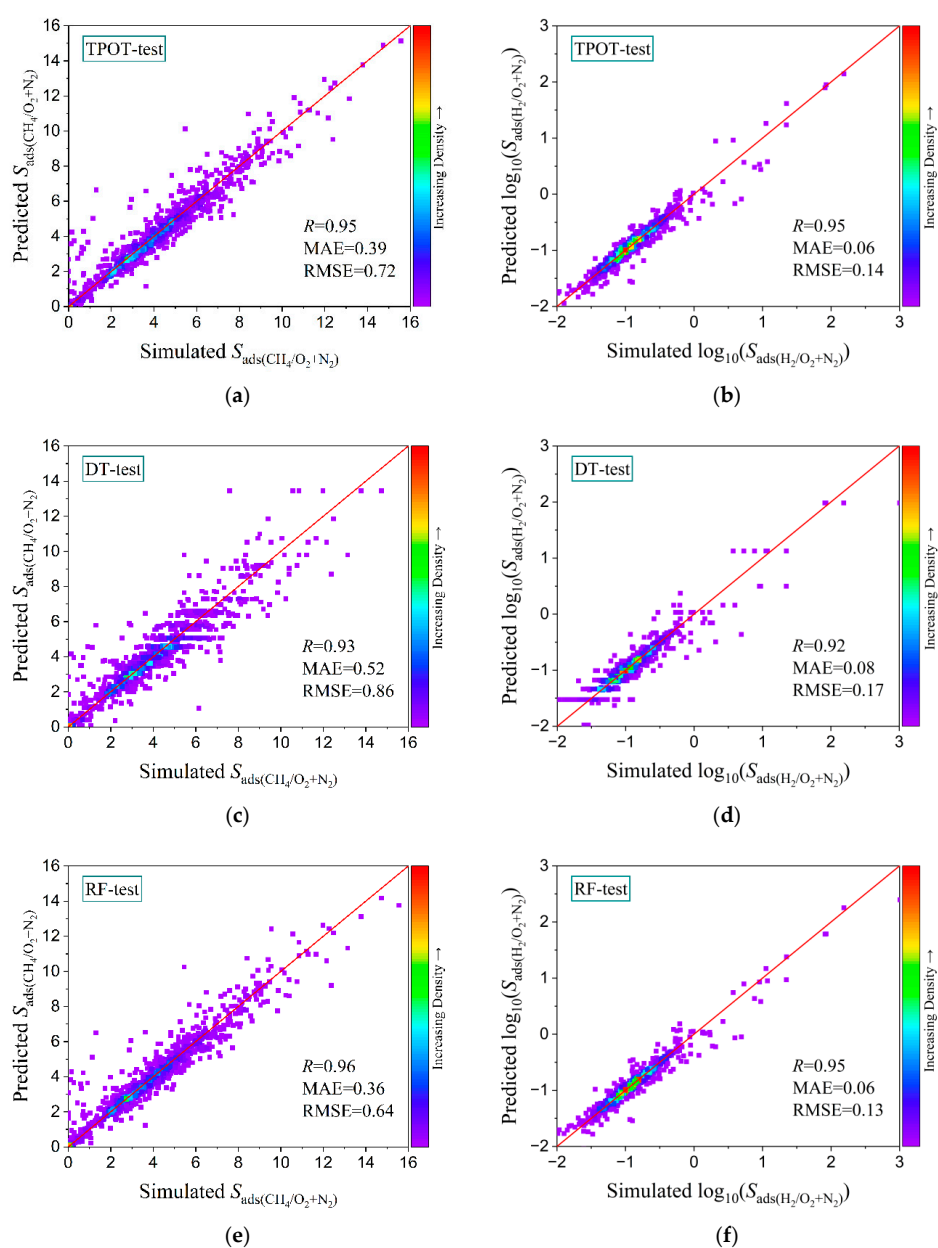


Figure 3. $S_{\text{ads(CH4/O2+N2)}}$ and $\log_{10}(S_{\text{ads(H2/O2+N2)})}$ predicted by three ML methods: (a,b) TPOT; (c,d) DT; and (e,f) RF versus the simulated values on the test set.

Table 1. Evaluation of three ML algorithms for the systems of CH₄/(O₂ + N₂) and H₂/(O₂ + N₂).

System	Performance Indicators	Machine Learning Methods	Training Set			Test Set		
			R	MAE	RMSE	R	MAE	RMSE
CH ₄ /(O ₂ + N ₂)	D_{CH_4}	TPOT	0.93	1.20	3.44	0.88	1.48	3.15
		DT	0.86	1.98	4.70	0.80	1.87	3.96
		RF	0.91	1.46	3.87	0.85	2.08	4.42
	$S_{ads(CH_4/O_2+N_2)}$	TPOT	0.98	0.21	0.56	0.95	0.39	0.72
		DT	0.96	0.39	0.72	0.93	0.52	0.86
		RF	0.99	0.15	0.42	0.96	0.36	0.64
	$S_{diff(CH_4/O_2+N_2)}$	TPOT	0.90	0.10	0.19	0.84	0.14	0.22
		DT	0.85	0.15	0.23	0.78	0.16	0.25
		RF	0.91	0.10	0.18	0.85	0.14	0.21
	$\log(P_{CH_4})$	TPOT	0.98	0.35	0.47	0.97	0.37	0.52
		DT	0.99	0.14	0.37	0.98	0.29	0.47
		RF	0.99	0.21	0.32	0.99	0.26	0.36
$\log(S_{perm(CH_4/O_2+N_2)})$	TPOT	0.97	0.19	0.34	0.96	0.21	0.38	
	DT	0.97	0.19	0.37	0.95	0.21	0.36	
	RF	0.98	0.14	0.29	0.97	0.17	0.31	
H ₂ /(O ₂ + N ₂)	D_{H_2}	TPOT	0.98	4.58	9.36	0.95	6.36	10.81
		DT	0.96	8.01	12.94	0.92	8.25	13.81
		RF	0.99	3.03	5.50	0.95	6.30	10.73
	$\log(S_{ads(H_2/O_2+N_2)})$	TPOT	0.97	0.03	0.10	0.95	0.06	0.14
		DT	0.97	0.06	0.11	0.92	0.08	0.17
		RF	0.98	0.03	0.08	0.95	0.06	0.13
$\log(S_{diff(H_2/O_2+N_2)})$	TPOT	0.76	0.12	0.18	0.71	0.13	0.19	
	DT	0.71	0.13	0.20	0.65	0.13	0.21	
	RF	0.78	0.12	0.17	0.73	0.12	0.19	
P_{H_2}	TPOT	0.98	556.09	1226.85	0.94	1179.56	2367.19	
	DT	0.97	961.07	1862.85	0.91	1497.22	2896.62	
	RF	0.98	694.80	1418.32	0.93	1245.29	2498.44	
$S_{perm(H_2/O_2+N_2)}$	TPOT	0.97	0.13	0.19	0.96	0.17	0.25	
	DT	0.97	0.13	0.20	0.94	0.19	0.29	
		RF	0.99	0.07	0.11	0.96	0.15	0.23

In conclusion, for the prediction of all the performance indicators, the RF prediction effect was generally better than that of the DT model for different gas mixture systems, which may have been because the RF model has a strong generalization ability. In terms of the adsorption performances of the two systems, the TPOT algorithm was better than the DT algorithm for most of the predictions. This may have been because TPOT could find the best combination of models and parameters by using a genetic algorithm, which led to better prediction results on different data sets. Although the prediction results of RF and TPOT were similar, the R and error (MAE and RMSE) values of most of the performance indicators showed that the prediction accuracy of the RF algorithm was slightly higher than that of TPOT. Therefore, RF was considered to be an optimal, stable, and useful algorithm for the MOF system in this work. In many previous studies, the RF algorithm also showed a good prediction ability for MOF systems [56–58].

Based on the discussion above, RF was further used to estimate the influences of the seven descriptors of the MOFs on the performances of the MOF adsorbents and membranes. The results are shown in Figure 4. The relative importance (RI) analysis showed that in the CH₄/N₂ + O₂ system (Figure 4a), the Henry's constant of $S_{ads(CH_4/O_2+N_2)}$ had the highest importance (about 49%, more than two times that of Q_{st}), and the RI order was $K_{CH_4} > Q_{st(CH_4)} > LCD > \phi > \rho > PLD \geq VSA$. This is consistent with the conclusion of Cai et al. [59] that the Henry's constant is the key descriptor of the trade-offs of C_1 , C_2 , and C_3 . In addition, the RF model quantitatively showed that PLD was the first important descriptor of the CH₄ diffusion coefficient and diffusion selectivity, which was a similar conclusion to that of Yang et al. [60]. For P_{CH_4} , the RI values of LCD and $Q_{st(CH_4)}$ of the MOFMs were large, with values of 33.50% and 28.14%, respectively. At the same time, the

RF model analysis showed that K_{CH_4} was the most important descriptor of S_{perm} , and LCD was the second most important. In the $\text{H}_2/\text{N}_2 + \text{O}_2$ system (Figure 4b), LCD was the most important influencing factor for both S_{ads} and S_{diff} , with an RI of about 40%; ϕ was the most important descriptor of P , with a large RI (about 5.6 times that of ρ), and the RI order was $\phi > \rho > Q_{\text{st}(\text{H}_2)} > \text{PLD} > \text{VSA} \geq \text{LCD} > K_{\text{H}_2}$. In addition, ϕ was an important factor for S_{perm} , and the change of PLD (RI of about 67%) could also affect S_{perm} . These results showed that the RF algorithm could accurately predict the performance indicators of the two systems by using the relevant descriptors of the MOFs and quantitatively determine the importance of the descriptors. This is conducive to screening candidate materials from a large number of MOFs to further guide experiments.

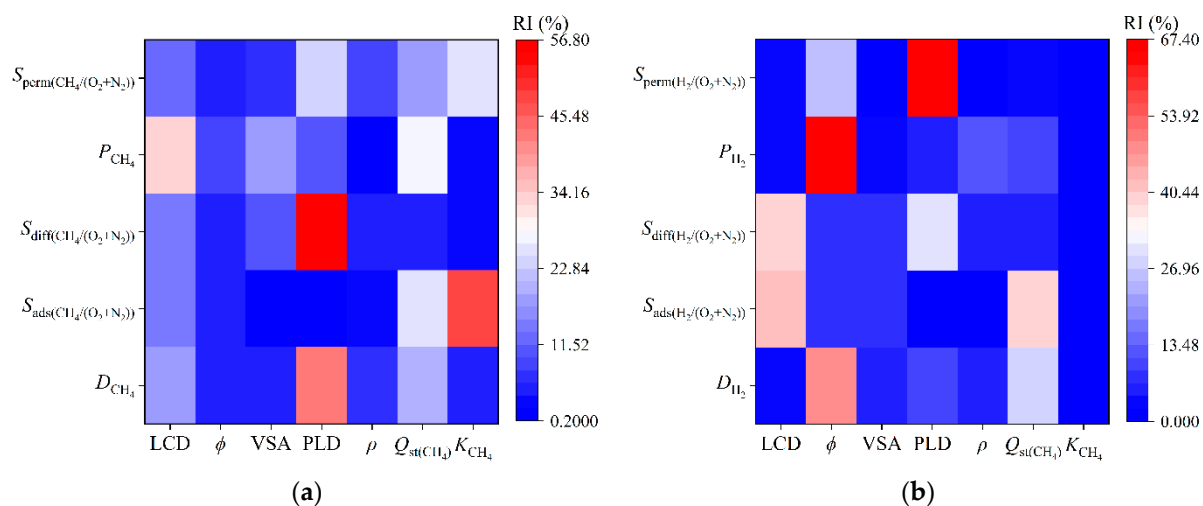


Figure 4. Relative importance values of the seven descriptors predicted by the RF algorithm for: (a) $\text{CH}_4/\text{O}_2 + \text{N}_2$; and (b) $\text{H}_2/\text{O}_2 + \text{N}_2$.

3.4. Top-Performing MOFs and MOFMs

In order to select MOFs and MOFMs with excellent performances for different systems, we used the limiting conditions shown in Table S4 to determine the best MOFs of the two systems as well as D , S_{ads} , S_{diff} , P , and S_{perm} of each adsorbate gas molecule in the MOFMs. Five optimal MOFs and five optimal MOFMs were screened for each system, which are listed in Table 2. In Figure 5a,b, the blue dots in the pink shaded part represent the highest-potential MOFs, which showed the best combination of adsorption and diffusion selectivity and had high D_{CH_4} or D_{H_2} values simultaneously. However, it was further found that most of the MOFs in Table S4 had large $S_{\text{diff}(\text{H}_2/(\text{O}_2+\text{N}_2))}$ values, but only when the diffusion coefficient of H_2 was larger, was the MOF more conducive to separation. Among the five materials selected for the $\text{CH}_4/\text{N}_2 + \text{O}_2$ mixed system, the $S_{\text{ads}(\text{CH}_4/\text{O}_2+\text{N}_2)}$, $S_{\text{diff}(\text{CH}_4/\text{O}_2+\text{N}_2)}$, and D_{CH_4} values of the ITAHEQ MOF were 7.61, 6.79, and $4.84 \times 10^{-6} \text{ cm}^2/\text{s}$, respectively, which were the largest adsorption and diffusion selectivity values of the materials considered. In addition, $S_{\text{diff}(\text{CH}_4/\text{O}_2+\text{N}_2)}$ was generally lower than $S_{\text{ads}(\text{CH}_4/\text{O}_2+\text{N}_2)}$, which also meant that the diffusion property of CH_4 was the key performance metric for the MOFs to capture low-concentration CH_4 from N_2 and O_2 during dynamic adsorption and membrane separation. Moreover, some MOFs in Figure 5a,c with very small pores only could adsorb single special gas molecules, leading to high selectivity. However, these MOFs also possess very low loading or permeability, so they are still not good candidates. This was consistent with the previous RI analysis. PLD is the key descriptor to control D_{CH_4} and $S_{\text{diff}(\text{CH}_4/\text{O}_2+\text{N}_2)}$. Likewise, Sumer [61] pointed out that CH_4 has a stronger adsorption capacity than N_2 , and in his work, the better performing MOFs screened showed high adsorption selectivity for CH_4 (1.3–9) over weak diffusivity selectivity for N_2 (1–4), which makes them methane-selective membranes. Qiao [26] also showed that diffusion is not only influenced by the size of gas molecules but also by gas–framework interactions, and

$S_{diff}(N_2/CH_4)$ is much larger than $S_{ads}(N_2/CH_4)$. The separation is controlled by diffusion, with methane diffusion being much slower and N_2/CH_4 separation being driven by diffusion. In the $H_2/N_2 + O_2$ system, the analysis of the optimal MOFs showed that the LCD range was very concentrated, ranging from 2.89 to 2.98 Å. This was because the N_2 and O_2 molecules in the LCD range could not enter the channel due to the large molecular dynamic diameters, which significantly improved the H_2 separation performance. At the same time, according to the analysis of the RI of the descriptors calculated by the RF algorithm, LCD is the key descriptor to control $H_2/N_2 + O_2$ separation. This showed that LCD with a kinetic diameter close to the H_2 diameter was the key for capturing H_2 molecules from the air and achieving a good separation effect, which provides effective theoretical guidance for the design and synthesis of new MOFs. The optimal MOFMs of the two systems are shown with blue dots in Figure 5c,d, with the highest P and S_{perm} values. These optimal MOFMs often have narrow channel structures, especially in $H_2/N_2 + O_2$. LCD was mainly concentrated in the range of 2.74–2.83 Å, and PLD was in the range of 2.44–2.52 Å. For the optimal MOFMs for CH_4 separation, the PLD also ranged from 2.85–3.49 Å, but the values were higher than those for the H_2 separation membranes because the size of the CH_4 molecule is larger than that of H_2 . Thus, CH_4 requires larger pores to pass through the membrane. Since VSA was calculated by GCMC using N_2 gas molecules as probes, when the pore diameter was smaller than the N_2 molecular diameter, VSA was displayed as 0. Therefore, the VSA of almost all the optimal MOFMs was 0. These MOFMs showed excellent separation performances, and their P_{H_2} values were between 6000 and 10,000 barrer. S_{perm} was also above 10, which is suitable for H_2 membrane separation applications.

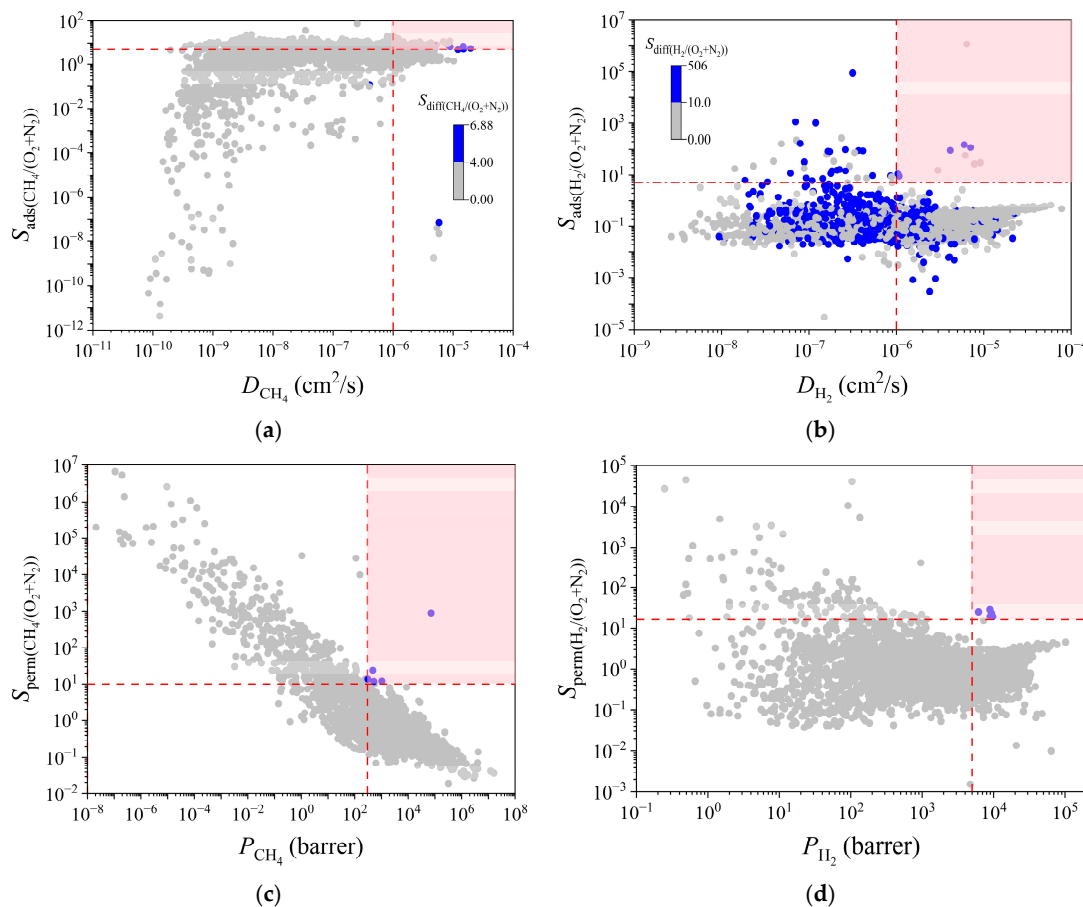


Figure 5. (a,b) Selectivity S_{ads} vs. diffusion coefficient D and (c,d) permeation selectivity S_{perm} vs. permeability P for $CH_4/O_2 + N_2$ and $H_2/O_2 + N_2$ in 6013 CoRE-MOFs. The color coding represents the diffusion selectivity S_{diff} .

Table 2. Best MOFs and MOFMs.

Application	Systems	CSD Code	LCD (Å)	ϕ	VSA (m ² /cm ³)	PLD (Å)	ρ (kg/m ³)	Q_{st} (kJ/mol)	K (mol/kg/Pa)	D_{CH_4} (10 ⁻⁶ cm ² /s)	S_{ads}	S_{diff}	
MOFs	CH ₄ /N ₂ + O ₂	ITAHEQ	4.67	0.15	105.10	4.13	1774.31	25.29	1.26×10^{-5}	4.84	7.61	6.79	
		QATLEE	4.25	0.31	289.95	4.04	2185.04	25.25	2.28×10^{-5}	19.58	5.57	6.88	
		XEJVOZ	5.05	0.16	274.91	4.70	2255.03	24.01	1.35×10^{-5}	14.76	7.05	4.93	
		FUDQIF	4.38	0.37	261.53	3.86	1573.76	27.47	7.58×10^{-5}	8.92	6.64	4.69	
		REGJIW	4.22	0.30	239.00	4.03	2185.50	24.82	1.73×10^{-5}	14.90	5.11	5.98	
	H ₂ /N ₂ + O ₂	ja4050828	2.89	0.04	0.00	2.41	1779.43	4.57	3.3×10^{-9}	5.97	153.95	504.05	
		ja4044642_si_002	2.90	0.04	0.00	2.43	1773.83	4.60	3.6×10^{-9}	7.08	119.50	87.59	
		ja403810k_si_003	2.92	0.04	0.00	2.44	1764.65	4.80	4.2×10^{-9}	4.14	91.54	161.42	
		UMEMAB	2.97	0.09	0.00	2.51	2543.92	6.19	8.8×10^{-9}	1.04	11.22	19.43	
		IFUDAO	2.98	0.12	0.00	2.84	1772.35	6.11	1.9×10^{-9}	1.08	9.03	17.63	
MOFMs	CH ₄ /N ₂ + O ₂	GOJRED	7.79	0.34	362.41	2.85	1251.40	40.28	4.47×10^{-4}	72177.49	892.96		
		ZIHTEP	4.35	0.17	108.53	3.49	1109.96	20.74	3.85×10^{-6}	475.23	23.97		
		XORGUI	4.07	0.35	58.28	3.22	1719.10	26.83	1.49×10^{-5}	312.12	13.53		
		LULJAE	3.65	0.30	57.51	3.28	1414.94	23.86	8.18×10^{-6}	1038.17	11.98		
		YAYPAR	3.81	0.18	4.86	3.12	2963.52	25.23	4.16×10^{-6}	521.57	11.72		
	H ₂ /N ₂ + O ₂	WENSIS	2.75	0.25	0.00	2.44	1625.34	16.23	9.07×10^{-7}	8900.57	28.15		
		IDAZEU	2.74	0.25	0.00	2.46	2245.15	16.40	6.73×10^{-7}	9354.21	22.65		
		IQUNAJ01	2.83	0.25	0.00	2.52	1719.74	16.28	8.85×10^{-7}	9726.40	19.04		
		IQUNAJ	2.83	0.24	0.00	2.52	1726.63	16.44	8.63×10^{-7}	8950.93	19.30		
		YAFGAP	2.75	0.24	0.00	2.44	2321.42	14.40	3.86×10^{-7}	6142.86	24.84		
												P (barrer) ^a	S_{perm}

^a 1 barrer = 3.348×10^{-16} mol·m/(m²·s·Pa) = 10^{-10} cm³ (STP) cm/(cm²·s·cmHg).

3.5. Design Strategies of MOFs and MOFMs with High Performances

In order to further guide the experiment and design and synthesize new MOF adsorbents or membranes with excellent performance, this section discusses a series of design strategies for a series of new hypothetical MOF materials based on ml and HTCS predictions. Firstly, 10 pairs of materials (20 pairs in total) are selected for each system. Each pair contains one MOF with excellent performance and one MOF with poor performance, but only one MOF composition has changed between them, such as topology, connector, or metal center, as shown in Table S5. Based on this, we propose three design strategies to facilitate the capture of CH₄ or H₂ from the air, as shown in Figure 6. Three pairs of representative design strategies are selected for each system, similar to the previous research work [34,62]. Among them, Figure 6a–c show three strategies for the CH₄/O₂ + N₂ system, and Figure 6d–f show three strategies for H₂/O₂ + N₂ system. Figure 6a,d display that changes in organic linkers result in different separation effects. Combined with RL, K_{CH_4} , and LCD, their adsorption selectivity and permeability were greatly affected. In Figure 6a for MOFMs, different organic connectors change the size of the channel, making the channel more compact. The LCD is reduced from 6.906 Å to 5.769 Å. Both LCD and PLD are closer to the dynamic diameter (3.8 Å) of CH₄, thus enhancing the separation effect of CH₄. For MOF adsorbent, the Henry's constant of MOF with excellent performance is relatively large ($1.36 \times 10^{-5} > 8.07 \times 10^{-6}$), plus with above influence of LCD, CH₄ can be adsorbed preferentially. Figure 6b,e have the same topology and organic links but different metal centers. In Figure 6b, MOF with good performance takes V as the metal center. Since the atomic radius is similar to that of Cr, but the metal activity of V is stronger than that of Cr, it has better adsorption, separation, or permeability. Figure 6e depicts that the atomic radius of metal Zn is smaller than that of Cd. Therefore, the regulation of the metal center affects the pore size of MOF and then affects the properties. S_{ads} , S_{diff} , and S_{perm} all increased by 60%~300%. Therefore, finding suitable metal nodes for a certain system can be considered an effective strategy to improve the separation performance. Figure 6c,f show the same metal centers and organic linkers, but different topologies lead to different separation effects. Among them, Figure 6c shows two different topological structures (mmm and rtl) assembled by the Cd-based metal center and 1,3,5-benzenetricarboxylic acid. Because the pore channels formed by different topological networks have different pore diameters and shapes, the diffusion coefficient is increased by 79%, and the gas permselectivity is increased by 120%. Figure 6f shows two different topologies (hcb and mog) assembled

by the Ag-based metal center and 1,3,5,7-tetrazatricyclo[3.3.1.1.3,7]decane. The pore shape is more complex and has better performance of H₂ separation. Therefore, for different systems, better MOF adsorbents or membranes can be obtained or designed by adjusting more appropriate topological structures, connectors, or metal centers.

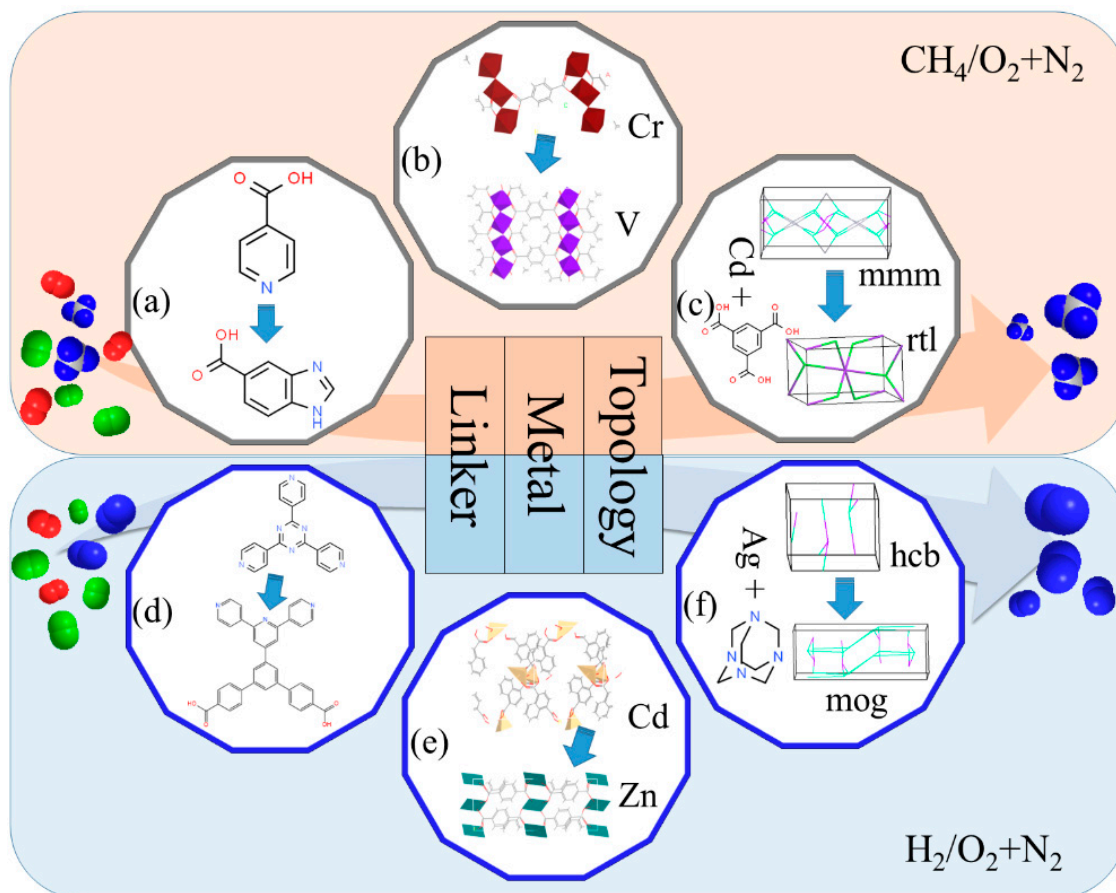


Figure 6. Three design strategies for boosting the separation of (a–c) CH₄/O₂ + N₂; and (d–f) H₂/O₂ + N₂.

4. Conclusions

In this work, the adsorption separation performances of 6013 CoRE-MOF adsorbents and membranes for CH₄/N₂ + O₂ and H₂/N₂ + O₂ mixtures were simulated using large-scale computational screening and ML. First, the close relationship between $D(P)$ and PLD in both systems was initially examined by univariate analysis, and the quantitative relationships $\log D = a\text{PLD} - b$ and $\log P = c\text{PLD} - d$ were established. Then, three ML methods, TOPT, DT, and RF, were used to predict the performance indices of each system. The analysis of the results showed that the model prediction accuracies were ranked as RF \geq TPOT > DT. Moreover, the feature importance was determined by the RF algorithm, and the RI ranking for the adsorption selectivity prediction of CH₄ was $K_{\text{CH}_4} > Q_{\text{st}(\text{CH}_4)} > \text{LCD} > \phi > \rho > \text{PLD} \geq \text{VSA}$. Nevertheless, the descriptor that had the greatest impact on D_{CH_4} and $S_{\text{diff}(\text{CH}_4/\text{O}_2+\text{N}_2)}$ was PLD. LCD and $Q_{\text{st}(\text{CH}_4)}$ were the two important descriptors of P and S_{perm} . In the H₂/N₂ + O₂ system, LCD had the largest RI for the performance of the H₂ adsorbent, while for P , the RI ranking was $\phi > \rho > Q_{\text{st}(\text{H}_2)} > \text{PLD} > \text{VSA} \geq \text{LCD} > K_{\text{H}_2}$. Furthermore, five optimal MOFs and five optimal MOFMs were screened for each of the two systems. The analysis of the optimal MOFs also further verified that the diffusivity of CH₄ is a key indicator to test in dynamic adsorption or membrane separation processes, and an LCD close to the kinetic diameter of H₂ is an overriding condition to achieve the separation of H₂ from air. Finally, three

types of design strategies, tuned topology, alternative metal nodes, and organic linkers, were proposed to effectively facilitate the capture of low-concentration CH₄ or H₂ from the air for specific applications. This study may provide an effective guide for experimental researchers to find MOFs/MOFMs for capturing energy gases (CH₄ and H₂) from air and can lead to new research ideas for applications involving coal mine spent air and hydrogen leakage.

Supplementary Materials: The supplementary material can be downloaded at: <https://www.mdpi.com/article/10.3390/membranes12090830/s1>. Figure S1: The relationship between the descriptors of MOFs with adsorbents performance in CH₄/N₂ + O₂; Figure S2: The relationship between the descriptors of MOFs with adsorbents performance in H₂/N₂ + O₂; Figure S3: The relationship between the descriptors of MOFMs with adsorbents performance in CH₄/N₂ + O₂; Figure S4: The relationship between the descriptors of MOFMs with adsorbents performance in H₂/N₂ + O₂; Figure S5: Diffusion coefficient D and Permeability P versus PLD for N₂ and O₂ in 6013 CORE-MOFs; Figure S6: Tree-based pipeline optimization tool; Figure S7: Decision tree; Figure S8: Random forest; Figure S9: k-fold cross validation; Table S1: Lennard–Jones parameters of MOFs; Table S2: Lennard–Jones parameters and charges of adsorbates; Table S3: Kinetic diameter of CH₄, N₂, O₂ and H₂; Table S4: The constraints for screening the best MOFs/MOFMs under each system; Table S5: Design Strategies of MOFs and MOFMs with high performance. References [48,63–66] are cited in the supplementary materials.

Author Contributions: Conceptualization, Z.Q., X.Z. and Y.Y.; methodology, Z.Q., Y.Y. and H.L.; software, Z.Q., X.Z. and Y.Y.; validation, H.L., Y.Z. and C.W.; formal analysis, D.L., Y.Z. and C.W.; investigation, X.Z.; resources, Z.Q.; data curation, Z.Q.; writing—original draft preparation, H.L., Y.Y. and Z.Q.; writing—review and editing, X.Z., Y.Y. and Z.Q.; visualization, H.L.; supervision, Y.Y.; project administration, X.Z.; funding acquisition, Z.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (21978058 and 21676094), the Pearl River Talent Recruitment Program (2019QN01L255), the Natural Science Foundation of Guangdong Province (2022A1515011446 and 2020A1515010800), Guangzhou Municipal Science and Technology Project (202102020875) and the R & D Program of Joint Institute of GZHU & ICoST (GI202102).

Acknowledgments: We gratefully thank the National Natural Science Foundation of China (21978058 and 21676094), the Pearl River Talent Recruitment Program (2019QN01L255), the Natural Science Foundation of Guangdong Province (2022A1515011446 and 2020A1515010800), Guangzhou Municipal Science and Technology Project (202102020875) and the R & D Program of Joint Institute of GZHU & ICoST (GI202102) for financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, W.; Xia, X.; Li, S. Large-scale evaluation of cascaded adsorption heat pumps based on metal/covalent–organic frameworks. *J. Mater. Chem. A* **2019**, *7*, 25010–25019. [[CrossRef](#)]
2. Lin, R.-B.; Xiang, S.; Zhou, W.; Chen, B. Microporous metal-organic framework materials for gas separation. *Chem* **2020**, *6*, 337–363. [[CrossRef](#)]
3. Van Vleet, M.J.; Weng, T.T.; Li, X.Y.; Schmidt, J.R. In situ, time-resolved, and mechanistic studies of metal-organic framework nucleation and growth. *Chem. Rev.* **2018**, *118*, 3681–3721. [[CrossRef](#)] [[PubMed](#)]
4. Lu, Y.; Zhang, H.; Chan, J.Y.; Ou, R.; Zhu, H.; Forsyth, M.; Marijanovic, E.M.; Doherty, C.M.; Marriott, P.J.; Holl, M.M.B.; et al. Homochiral MOF–polymer mixed matrix membranes for efficient separation of chiral molecules. *Angew. Chem. Int. Ed.* **2019**, *58*, 16928–16935. [[CrossRef](#)] [[PubMed](#)]
5. Wang, M.; Wang, T.; Cai, P.; Chen, X. Nanomaterials discovery and design through machine learning. *Small Methods* **2019**, *3*, 1900025. [[CrossRef](#)]
6. Shi, Z.; Liang, H.; Yang, W.; Liu, J.; Liu, Z.; Qiao, Z. Machine learning and in silico discovery of metal-organic frameworks: Methanol as a working fluid in adsorption-driven heat pumps and chillers. *Chem. Eng. Sci.* **2020**, *214*, 115430. [[CrossRef](#)]
7. Yuan, X.; Li, L.; Shi, Z.; Liang, H.; Li, S.; Qiao, Z. Molecular-fingerprint machine-learning-assisted design and prediction for high-performance MOFs for capture of NMHCs from air. *Adv. Powder Mater.* **2022**, *1*, 100026. [[CrossRef](#)]
8. He, T.; Pachfule, P.; Wu, H.; Xu, Q.; Chen, P. Hydrogen carriers. *Nat. Rev. Mater.* **2016**, *1*, 16059. [[CrossRef](#)]

9. Feng, L.; Wang, Y.; Zhang, K.; Wang, K.-Y.; Fan, W.; Wang, X.; Powell, J.A.; Guo, B.; Dai, F.; Zhang, L.; et al. Molecular pivot-hinge installation to evolve topology in rare-earth metal-organic frameworks. *Angew. Chem. Int. Ed.* **2019**, *58*, 16682–16690. [[CrossRef](#)]
10. Luo, L.; Lo, W.S.; Si, X.; Li, H.; Wu, Y.; An, Y.; Zhu, Q.; Chou, L.Y.; Li, T.; Tsung, C.K. Directional engraving within single crystalline metal-organic framework particles via oxidative linker cleaving. *J. Am. Chem. Soc.* **2019**, *141*, 20365–20370. [[CrossRef](#)]
11. Zhang, F.; Liu, Y.; Lei, J.; Wang, S.; Ji, X.; Liu, H.; Yang, Q. Metal-organic-framework-derived carbon nanostructures for site-specific dual-modality photothermal/photodynamic thrombus therapy. *Adv. Sci.* **2019**, *6*, 1901378. [[CrossRef](#)] [[PubMed](#)]
12. Liu, D.; Lu, K.; Poon, C.; Lin, W. Metal-organic frameworks as sensory materials and imaging agents. *Inorg. Chem.* **2014**, *53*, 1916–1924. [[CrossRef](#)] [[PubMed](#)]
13. Chang, M.; Ren, J.; Yang, Q.; Liu, D. A robust calcium-based microporous metal-organic framework for efficient CH₄/N₂ separation. *Chem. Eng. J.* **2021**, *408*, 127294. [[CrossRef](#)]
14. Xu, G.J.; Meng, Z.S.; Liu, Y.Z.; Guo, X.J.; Deng, K.M.; Ding, L.F.; Lu, R.F. Porous MOF-205 with multiple modifications for efficiently storing hydrogen and methane as well as separating carbon dioxide from hydrogen and methane. *Int. J. Energ. Res.* **2019**, *43*, 7517–7528. [[CrossRef](#)]
15. Kang, Z.; Xue, M.; Fan, L.; Huang, L.; Guo, L.; Wei, G.; Chen, B.; Qiu, S. Highly selective sieving of small gas molecules by using an ultra-microporous metal-organic framework membrane. *Energ. Environ. Sci.* **2014**, *7*, 4053–4060. [[CrossRef](#)]
16. Hou, Q.; Wu, Y.; Zhou, S.; Wei, Y.; Caro, J.; Wang, H. Ultra-tuning of the aperture size in stiffened ZIF-8_{Cm} frameworks with mixed-linker strategy for enhanced CO₂/CH₄ separation. *Angew. Chem. Int. Ed.* **2019**, *58*, 327–331. [[CrossRef](#)]
17. Gulbalkan, H.C.; Haslak, Z.P.; Altintas, C.; Uzun, A.; Keskin, S. Assessing CH₄/N₂ separation potential of MOFs, COFs, IL/MOF, MOF/Polymer, and COF/Polymer composites. *Chem. Eng. J.* **2022**, *428*, 131239. [[CrossRef](#)]
18. Belmabkhout, Y.; Mouttaki, H.; Eubank, J.F.; Guillerme, V.; Eddaoudi, M. Effect of pendant isophthalic acid moieties on the adsorption properties of light hydrocarbons in HKUST-1-like tbo-MOFs: Application to methane purification and storage. *RSC Adv.* **2014**, *4*, 63855–63859. [[CrossRef](#)]
19. Kang, Z.; Fan, L.; Sun, D. Recent advances and challenges of metal-organic framework membranes for gas separation. *J. Mater. Chem. A* **2017**, *5*, 10073–10091. [[CrossRef](#)]
20. Fan, H.; Peng, M.; Strauss, I.; Mundstock, A.; Meng, H.; Caro, J. MOF-in-COF molecular sieving membrane for selective hydrogen separation. *Nat. Commun.* **2021**, *12*, 38. [[CrossRef](#)]
21. Yang, L.; Qian, S.; Wang, X.; Cui, X.; Chen, B.; Xing, H. Energy-efficient separation alternatives: Metal-organic frameworks and membranes for hydrocarbon separation. *Chem. Soc. Rev.* **2020**, *49*, 5359–5406. [[CrossRef](#)] [[PubMed](#)]
22. Zhang, Y.; Feng, X.; Yuan, S.; Zhou, J.; Wang, B. Challenges and recent advances in MOF-polymer composite membranes for gas separation. *Inorg. Chem. Front.* **2016**, *3*, 896–909. [[CrossRef](#)]
23. Liu, G.; Chernikova, V.; Liu, Y.; Zhang, K.; Belmabkhout, Y.; Shekhah, O.; Zhang, C.; Yi, S.; Eddaoudi, M.; Koros, W.J. Mixed matrix formulations with MOF molecular sieving for key energy-intensive separations. *Nat. Mater.* **2018**, *17*, 283. [[CrossRef](#)] [[PubMed](#)]
24. Wang, Y.; Fin, H.; Ma, Q.; Mo, K.; Mao, H.; Feldhoff, A.; Cao, X.; Li, Y.; Pan, F.; Jiang, Z. A MOF glass membrane for gas separation. *Angew. Chem. Int. Ed.* **2020**, *59*, 4365–4369. [[CrossRef](#)]
25. Watanabe, T.; Sholl, D.S. Accelerating applications of metal-organic frameworks for gas adsorption and separation by computational screening of materials. *Langmuir* **2012**, *28*, 14114–14128. [[CrossRef](#)]
26. Qiao, Z.; Peng, C.; Zhou, J.; Jiang, J. High-throughput computational screening of 137953 metal-organic frameworks for membrane separation of a CO₂/N₂/CH₄ mixture. *J. Mater. Chem. A* **2016**, *4*, 15904–15912. [[CrossRef](#)]
27. Qiao, Z.; Xu, Q.; Jiang, J. Computational screening of hydrophobic metal-organic frameworks for the separation of H₂S and CO₂ from natural gas. *J. Mater. Chem. A* **2018**, *6*, 18898–18905. [[CrossRef](#)]
28. McIntyre, S.M.; Shan, B.; Wang, R.; Zhong, C.; Liu, J.; Mu, B. Monte carlo simulations to examine the role of pore structure on ambient air separation in metal-organic frameworks. *Ind. Eng. Chem. Res.* **2018**, *57*, 9240–9253. [[CrossRef](#)]
29. Qiao, Z.; Li, L.; Li, S.; Liang, H.; Zhou, J.; Snurr, R.Q. Molecular fingerprint and machine learning to accelerate design of high-performance homochiral metal-organic frameworks. *AIChE J.* **2021**, *67*, e17352. [[CrossRef](#)]
30. Shi, Z.; Yuan, X.; Yan, Y.; Tang, Y.; Li, J.; Liang, H.; Tong, L.; Qiao, Z. Techno-economic analysis of metal-organic frameworks for adsorption heat pumps/chillers: From directional computational screening, machine learning to experiment. *J. Mater. Chem. A* **2021**, *9*, 7656–7666. [[CrossRef](#)]
31. Anderson, R.; Rodgers, J.; Argueta, E.; Biong, A.; Gómez-Gualdrón, D.A. Role of pore chemistry and topology in the CO₂ capture capabilities of MOFs: From molecular simulation to machine learning. *Chem. Mater.* **2018**, *30*, 6325–6337. [[CrossRef](#)]
32. Moosavi, S.M.; Nandy, A.; Jablonka, K.M.; Ongari, D.; Janet, J.P.; Boyd, P.G.; Lee, Y.; Smit, B.; Kulik, H.J. Understanding the diversity of the metal-organic framework ecosystem. *Nat. Commun.* **2020**, *11*, 4068. [[CrossRef](#)] [[PubMed](#)]
33. Li, H.; Li, L.; Lin, R.-B.; Zhou, W.; Zhang, Z.; Xiang, S.; Chen, B. Porous metal-organic frameworks for gas storage and separation: Status and challenges. *EnergyChem* **2019**, *1*, 100006. [[CrossRef](#)]
34. Yan, Y.; Shi, Z.; Li, H.; Li, L.; Yang, X.; Li, S.; Liang, H.; Qiao, Z. Machine learning and in-silico screening of metal-organic frameworks for O₂/N₂ dynamic adsorption and separation. *Chem. Eng. J.* **2022**, *427*, 131604. [[CrossRef](#)]
35. Tang, H.; Xu, Q.; Wang, M.; Jiang, J. Rapid screening of metal-organic frameworks for propane/propylene separation by synergizing molecular simulation and machine learning. *ACS Appl. Mater. Inter.* **2021**, *13*, 53454–53467. [[CrossRef](#)]

36. Rosen, A.S.; Iyer, S.M.; Ray, D.; Yao, Z.; Aspuru-Guzik, A.; Gagliardi, L.; Notestein, J.M.; Snurr, R.Q. Machine learning the quantum-chemical properties of metal-organic frameworks for accelerated materials discovery. *Matter* **2021**, *4*, 1578–1597. [[CrossRef](#)]
37. Moosavi, S.M.; Chidambaram, A.; Talirz, L.; Haranczyk, M.; Stylianou, K.C.; Smit, B. Capturing chemical intuition in synthesis of metal-organic frameworks. *Nat. Commun.* **2019**, *10*, 539. [[CrossRef](#)]
38. Azar, A.N.V.; Velioglu, S.; Keskin, S. Large-scale computational screening of metal organic framework (MOF) membranes and MOF-based polymer membranes for H₂/N₂ separations. *ACS Sustain. Chem. Eng.* **2019**, *7*, 9525–9536. [[CrossRef](#)]
39. Qiao, Z.; Xu, Q.; Jiang, J. High-throughput computational screening of metal-organic framework membranes for upgrading of natural gas. *J. Membr. Sci.* **2018**, *551*, 47–54. [[CrossRef](#)]
40. Bai, X.; Shi, Z.; Xia, H.; Li, S.; Liu, Z.; Liang, H.; Liu, Z.; Wang, B.; Qiao, Z. Machine-Learning-Assisted High-Throughput computational screening of Metal-Organic framework membranes for hydrogen separation. *Chem. Eng. J.* **2022**, *446*, 136783. [[CrossRef](#)]
41. Skoulidas, A.I.; Sholl, D.S. Self-diffusion and transport diffusion of light gases in metal-organic framework materials assessed using molecular dynamics simulations. *J. Phys. Chem. B* **2005**, *109*, 15760–15768. [[CrossRef](#)] [[PubMed](#)]
42. Erucar, I.; Keskin, S. Computational assessment of MOF membranes for CH₄/H₂ separations. *J. Membr. Sci.* **2016**, *514*, 313–321. [[CrossRef](#)]
43. Adatoz, E.; Avci, A.K.; Keskin, S. Opportunities and challenges of MOF-based membranes in gas separations. *Sep. Purif. Technol.* **2015**, *152*, 207–237. [[CrossRef](#)]
44. Avci, G.; Velioglu, S.; Keskin, S. High-throughput screening of MOF adsorbents and membranes for H₂ purification and CO₂ capture. *ACS Appl. Mater. Inter.* **2018**, *10*, 33693–33706. [[CrossRef](#)] [[PubMed](#)]
45. Chung, Y.G.; Camp, J.; Haranczyk, M.; Sikora, B.J.; Bury, W.; Krungleviciute, V.; Yildirim, T.; Farha, O.K.; Sholl, D.S.; Snurr, R.Q. Computation-ready, experimental metal-organic frameworks: A tool to enable high-throughput screening of nanoporous crystals. *Chem. Mater.* **2014**, *26*, 6185–6192. [[CrossRef](#)]
46. Willems, T.F.; Rycroft, C.H.; Kazi, M.; Meza, J.C.; Haranczyk, M. Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Micropor. Mesopor. Mat.* **2012**, *149*, 134–141. [[CrossRef](#)]
47. Zhou, Y.P.; Wei, L.F.; Yang, J.; Sun, Y.; Zhou, L. Adsorption of oxygen on superactivated carbon. *J. Chem. Eng. Data* **2005**, *50*, 1068–1072. [[CrossRef](#)]
48. Potoff, J.J.; Siepmann, J.I. Vapor-liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen. *AIChE J.* **2001**, *47*, 1676–1682. [[CrossRef](#)]
49. Garberoglio, G.; Skoulidas, A.I.; Johnson, J.K. Adsorption of gases in metal organic materials: Comparison of simulations and experiments. *J. Phys. Chem. B* **2005**, *109*, 13094–13103. [[CrossRef](#)]
50. Qiao, Z.; Xu, Q.; Cheetham, A.K.; Jiang, J. High-throughput computational screening of metal-organic frameworks for thiol capture. *J. Phys. Chem. C* **2017**, *121*, 22208–22215. [[CrossRef](#)]
51. Ewald, P.P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369*, 253–287. [[CrossRef](#)]
52. Hantal, G.; Jedlovsky, P.; Hoang, P.N.M.; Picaud, S. Calculation of the adsorption isotherm of formaldehyde on ice by grand canonical Monte Carlo simulation. *J. Phys. Chem. C* **2007**, *111*, 14170–14178. [[CrossRef](#)]
53. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
54. Wilmer, C.E.; Farha, O.K.; Bae, Y.-S.; Hupp, J.T.; Snurr, R.Q. Structure-property relationships of porous materials for carbon dioxide separation and capture. *Energ. Environ. Sci.* **2012**, *5*, 9849–9856. [[CrossRef](#)]
55. Keskin, S.; Sholl, D.S. Efficient methods for screening of metal organic framework membranes for gas separations using atomically detailed models. *Langmuir* **2009**, *25*, 11786–11795. [[CrossRef](#)]
56. Yuan, X.; Deng, X.; Cai, C.; Shi, Z.; Liang, H.; Li, S.; Qiao, Z. Machine learning and high-throughput computational screening of hydrophobic metal-organic frameworks for capture of formaldehyde from air. *Green Energy Environ.* **2021**, *6*, 759–770. [[CrossRef](#)]
57. Deng, X.; Yang, W.; Li, S.; Liang, H.; Shi, Z.; Qiao, Z. Large-scale screening and machine learning to predict the computation-ready, experimental metal-organic frameworks for CO₂ capture from air. *Appl. Sci.* **2020**, *10*, 569. [[CrossRef](#)]
58. Wu, X.; Xiang, S.; Su, J.; Cai, W. Understanding quantitative relationship between methane storage capacities and characteristic properties of metal-organic frameworks based on machine learning. *J. Phys. Chem. C* **2019**, *123*, 8550–8559. [[CrossRef](#)]
59. Cai, C.; Li, L.; Deng, X.; Li, S.; Liang, H.; Qiao, Z. Machine learning and high-throughput computational screening of metal-organic framework for separation of methane/ethane/propane. *Acta Chim. Sin.* **2020**, *78*, 427–436. [[CrossRef](#)]
60. Yang, W.; Liang, H.; Peng, F.; Liu, Z.; Liu, J.; Qiao, Z. Computational screening of metal-organic framework membranes for the separation of 15 gas mixtures. *Nanomaterials* **2019**, *9*, 467. [[CrossRef](#)]
61. Sumer, Z.; Keskin, S. Adsorption- and membrane-based CH₄/N₂ separation performances of MOFs. *Ind. Eng. Chem. Res.* **2017**, *56*, 8713–8722. [[CrossRef](#)]
62. Tang, H.; Jiang, J. In silico screening and design strategies of ethane-selective metal-organic frameworks for ethane/ethylene separation. *AIChE J.* **2021**, *67*, e17025. [[CrossRef](#)]
63. Rappé, A.K.; Casewit, C.J.; Colwell, K.S.; Goddard, W.A., III; Skiff, W.M. Uff a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035. [[CrossRef](#)]

64. Stoll, J.; Vrabec, J.; Hasse, H. Vapor-liquid equilibria of mixtures containing nitrogen, oxygen, carbon dioxide, and ethane. *AIChE J.* **2003**, *49*, 2187–2198. [[CrossRef](#)]
65. Martin, M.G.; Siepmann, J.I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *J. Phys. Chem. B* **1998**, *102*, 2569–2577. [[CrossRef](#)]
66. Shah, M.S.; Tsapatsis, M.; Siepmann, J.I. Development of the Transferable Potentials for Phase Equilibria Model for Hydrogen Sulfide. *J. Phys. Chem. B* **2015**, *119*, 7041–7052. [[CrossRef](#)]