

Article

Large-Scale Screening and Machine Learning to Predict the Computation-Ready, Experimental Metal-Organic Frameworks for CO₂ Capture from Air

Xiaomei Deng, Wenyuan Yang, Shuhua Li, Hong Liang *, Zenan Shi * and Zhiwei Qiao *

Guangzhou Key Laboratory for New Energy and Green Catalysis, School of Chemistry and Chemical Engineering, Guangzhou University, Guangzhou 510006, China; 1705200045@e.gzhu.edu.cn (X.D.); 2111705055@e.gzhu.edu.cn (W.Y.); lish@gzhu.edu.cn (S.L.)

* Correspondence: lhong@gzhu.edu.cn (H.L.); zenanshi@126.com (Z.S.); zqiao@gzhu.edu.cn (Z.Q.)

Received: 19 December 2019; Accepted: 10 January 2020; Published: 13 January 2020



Abstract: The rising level of CO₂ in the atmosphere has attracted attention in recent years. The technique of capturing CO₂ from higher CO₂ concentrations, such as power plants, has been widely studied, but capturing lower concentrations of CO₂ directly from the air remains a challenge. This study uses high-throughput computer (Monte Carlo and molecular dynamics simulation) and machine learning (ML) to study 6013 computation-ready, experimental metal-organic frameworks (CoRE-MOFs) for CO₂ adsorption and diffusion properties in the air with very low concentrations of CO₂. First, the law influencing CO₂ adsorption and diffusion in air is obtained as a structure-performance relationship, and then the law influencing the performance of CO₂ adsorption and diffusion in air is further explored by four ML algorithms. Random forest (RF) was considered the optimal algorithm for prediction of CO₂ selectivity, with an *R* value of 0.981, and this algorithm was further applied to analyze the relative importance of each metal-organic framework (MOF) descriptor quantitatively. Finally, 14 MOFs with the best properties were successfully screened out, and it was found that a key to capturing a low concentration CO₂ from the air was the diffusion performance of CO₂ in MOFs. When the pore-limiting diameter (PLD) of a MOF was closer to the CO₂ dynamic diameter, this MOF could possess higher CO₂ diffusion separation selectivity. This study could provide valuable guidance for the synthesis of new MOFs in experiments that capture directly low concentration CO₂ from the air.

Keywords: CO₂ capture; Monte Carlo; machine learning; metal-organic framework; adsorption; diffusion

1. Introduction

It is well known that the amount of CO₂ discharged into the atmosphere increases with the rapid development of industry and population growth. In addition, deforestation, the large amount of CO₂ and other gases generated by the burning of fossil fuels such as coal, oil, and natural gas directly discharged into the atmosphere, and the emission of limestone roasting to produce cement have resulted in global carbon dioxide emissions increasing by 3.8% [1]. All of the above factors have aggravated carbon dioxide emissions, thereby increasing the urgency of counteracting the greenhouse effect and its associated global warming. The Kyoto Protocol and the Paris Agreement aim to control greenhouse gas emissions under the United Nations Framework Convention on Climate Change (UNFCCC), in which CO₂ is listed as a major greenhouse gas that needs to be mitigated or recycled [2]. The greenhouse gases include more than CO₂, however; in fact, the global warming potentials of CH₄ and N₂O are 25 times and 298 times that of CO₂, respectively. Nevertheless, due to its relatively large emission levels, CO₂ accounts for approximately 55% of the total greenhouse gas contribution [3,4].

Thus, it is obvious that the adsorption and separation of carbon dioxide from the air is particularly important. In addition, the successful capture of CO₂ could have multifaceted practical values: first, oil recovery could be improved through appropriate reservoir engineering; second, the captured CO₂ could be used to produce industrial chemicals, including concrete, paint, and fertilizer; third, the CO₂ in the atmosphere could be captured and combined with hydrogen for direct synthesis into liquid hydrocarbons, which could then be utilized in fuel synthesis and supply, including gasoline and diesel. The use of raw materials can reduce the proportion of fossil energy to further control CO₂ emissions, ultimately achieving carbon neutrality or even net negative carbon emissions [5].

Recently, carbon engineering has developed a series of capture technologies that remove carbon dioxide directly from the air. Carbon dioxide can be removed from the atmosphere using biological, chemical, or physical processes [6]. These methods have certain limitations, however. For example, biological processes are very economical, but they are usually very slow and ineffective. As for chemical processes, the waste of carbon resources and volatilization of organic solvents during these actions lead to further environmental pollution, equipment corrosion, and complex post-treatment issues. The traditional technique for separating carbon dioxide is solvent washing, such as the use of an alcohol amine solution [7–10]. Although this conventional method can reduce the concentration of carbon dioxide in the air, it is extremely expensive, the solvent is difficult to regenerate, the operation is complicated, and it consumes a great deal of energy [11]. In fact, the energy consumption of solvent washing is 3 to 4 times that of CO₂ captured from exhaust gas [12]. Given these drawbacks, there is an urgent need to find a more efficient, convenient, and energy-saving technique to replace the traditional carbon dioxide capture method. Adsorption separation is a potential technique. It is not only inexpensive, but also simple in terms of operation and equipment, and relatively low in energy consumption when the adsorbent is regenerated (the regeneration process of adsorbents is to desorb the adsorbed substances). Conventional adsorbents, however, including activated carbon, zeolite, silica gel, and metal oxides have poor scavenging effects on carbon dioxide in the air due to inferior separation selectivity and regeneration difficulty. For example, silica gel, which has amorphous properties, does not have a continuous uniform porous structure and exhibits unfavorable diffusion properties [11]. Therefore, the development of a new type of adsorbent is imperative. In recent years, studies have shown that the use of metal-organic frameworks (MOFs) to adsorb and separate carbon dioxide can not only make up for the shortcomings of the above adsorbents, but also feature the advantages of high selectivity and being non-polluting. The MOF is an organic–inorganic hybrid material with intramolecular pores formed by the self-assembly of organic ligands and inorganic metal ions or clusters by coordination bonds [13]. Compared with common adsorbents, MOFs exhibit many advantages such as various structures and properties, large specific surface area, high porosity, and structural control. Therefore, they are widely used in gas adsorption [11] and separation [14–19], as well as general materials in processes including storage [20], optics [21], catalysis [22–25], and drug delivery [26,27]. To date, thousands of MOFs have been synthesized, some of which have been utilized in the attempt to capture CO₂ from the air. Peng et al. [28] designed and synthesized 2 incorporated MOFs to study their stability and ability to capture CO₂ from the air. Liu et al. [11] used an amine-functionalized MOF and an ultra-microporous MOF to capture CO₂ directly from the air, and further investigated the performance of CO₂ capture and the reproducibility of MOFs under humid conditions. Osama et al. [29] synthesized an isomorphic MOF SIFSIX-3-Cu with uniform adsorption sites for capturing CO₂ from the air. Since CO₂ capture from the air has a very high selectivity of MOF, when the traditional approach is used to screen MOFs for the best-performing candidates, it not only consumes a great deal of manpower and material resources, but also has an extended study period and causes pollution to a certain extent. With the continuous advancement and development of computers, molecular simulation is playing an increasingly important role in the field of materials science [30]. Some studies have used high-throughput molecular simulation calculation methods to screen large numbers of MOFs in a database, thereby successfully screening MOFs with high selectivity and high working capacity based on different target performance requirements. For

example, Wilmer et al. adsorbed pure carbon dioxide, nitrogen, and methane using more than 130,000 hypothetical MOFs, and proposed a relationship between structural properties (pore size, volume, and surface area) and chemical functions, as well as evaluation criteria for the separation of carbon dioxide from adsorbents [31]. In the presence of nickel dilution, Watanabe et al. combined pore size analysis with classical simulation to screen 1163 MOFs as membrane materials for CO₂/N₂ separation [32]. Lin et al. screened hundreds of thousands of theoretically predicted zeolites and zeolite MOFs and identified a number of potential materials for capturing carbon dioxide [33]. Based on 105 MOFs, Wu et al. proposed the relationship of CO₂/N₂ adsorption selectivity with porosity and the isosteric heat of adsorption [34]. Fernandez et al. [35] used advanced machine-learning (ML) algorithms to quickly identify 292,050 hypothetical high-performance MOFs for pure CO₂ adsorption (0.15 bar and 1 bar). These screening studies, however, were aimed at capturing high concentrations of CO₂. Given that the concentration of CO₂ in the atmosphere is comparatively low relative to the concentrations of natural gas and other components, it is undoubtedly a challenge to discover efficient MOF materials that can directly capture CO₂ from the air.

To date, given that there have been 6013 MOFs reported, finding the appropriate MOFs for a specific system in such a large database is undoubtedly a daunting task. This study focused on the aforementioned MOF simulation of the adsorption and diffusion performances of CO₂, N₂, and O₂ in infinite dilutions in order to identify materials with excellent performance in terms of both static adsorption and kinetic adsorption. The influencing factors affecting the adsorption and diffusion of CO₂ were obtained by univariate analysis. Next, multivariate analyses, namely 4 ML algorithms (back propagation neural network (BPNN), decision tree (DT), random forest (RF), and support vector machine (SVM)), were explored in depth. Finally, we adopted the optimal algorithm model. The parameters affecting CO₂ selectivity were predicted, and 14 types of MOFs with the same diffusion selectivity and adsorption selectivity were selected.

2. Materials and Methods

2.1. Molecular Model

In this work, we used molecular simulation to screen the capability of 6013 computation-ready, experimental MOFs (CORE-MOFs version 2) [36] to capture CO₂ from the air. Their crystal structures were derived from the Cambridge Crystallographic Data Centre (CCDC), and their parameters were compiled and verified by Chung et al. [37]. We removed all solvent and ligand molecules prior to running the simulation. Each MOF used 5 structural parameters, namely, volumetric surface area (VSA), largest cavity diameter (LCD), pore-limiting diameter (PLD), porosity ϕ , density ρ , and an energy parameter: heat of adsorption. Both LCD and PLD were calculated using the Zeo++ software package [38]. The VSA and ϕ were calculated using the N₂ of 3.64 Å and He of 2.58 Å as probes in the RASPA software package [39]. If VSA is close to or equal to 0, this indicates that the MOF cannot accommodate N₂ molecules [40]. We used NVT-Monte Carlo (NVT-MC) simulation, where N is the number of particles, V is the volume of the system, and T is the temperature of the system. The Q_{st} of each gas was calculated in an infinite dilution state.

The force field parameters for the 3 gas components CO₂/N₂/O₂ were from the transferable potentials for phase equilibria (TraPPE) force field [41] and are listed in Table S2. The CO₂ molecule has a C-O bond length of 1.16 Å and a bond angle \angle OCO of 180°. N₂ is considered as a 3-point model, and the bond length of N-N is 1.10 Å. Oxygen is also a 3-point atom, and the O-O bond length is 1.21 Å. The models of 3 gases are shown in Figure S1. The atomic charge of MOF was estimated using the MOF electrostatic-potential-optimized charge scheme (MEPO-Qeq) method [42], which accurately evaluated electrostatic interactions. Due to the advantages of the MEPO-Qeq method with fast and accurate, it is widely used in various systems of adsorption-MOF [43–45]. The Lennard–Jones (LJ) electrostatic parameters were obtained from the universal force field (UFF) [46] and are listed in Table S1. Data from previous studies had shown that the UFF–TraPPE force field combination could accurately predict

the adsorption and diffusion behaviors of these 3 gases in MOFs [40,47,48]. The Lorentz–Berthelot combination rule was used to calculate the cross-LJ parameters.

2.2. Screening Methods

In MOFs, the values of Henry's constant K and the diffusion coefficient D of CO_2 , N_2 , and O_2 were estimated using Monte Carlo (MC) and molecular dynamics (MD) simulations with the same set, respectively. In principle, a single gas molecule should be added to an MOF to simulate infinite dilution, while in reality, we added 30 gas molecular models to each MOF, ignoring the force between the gas models, thus being equivalent to the independent simulation of each gas molecule. Ultimately, the simulation results of the 30 independent molecules were statistically averaged. Throughout the simulation, the MOF frame was assumed to be rigid and the simulation elements were extended to at least 24 Å along the three-dimensional periodic boundary conditions. A 12 Å spherical cutoff with long-range correction was used to calculate the LJ interaction, while the Ewald sum was used to calculate the electrostatic interaction. In each MOF, the MC simulation ran 100,000 cycles, with the first 50,000 used for balancing and the last 50,000 used for overall averaging. Each cycle consisted of n trials (n : number of adsorbed molecules), including translation, rotation, regeneration, and exchange (exchange movement, including insertion and deletion). In the MD simulation, the 30 gas molecules had an MD duration of 10 ns at each MOF, and 5 ns was ultimately selected for statistical averaging. After the sampling analysis of dozens of MOFs, it was found that further increases of cycle time and MD duration had little effect on the simulation results. All MCs and MDs were simulated using the RASPA software package [39].

3. Results and Discussion

3.1. Univariate Analysis

In order to investigate the relationship of CO_2 adsorption and diffusion properties in N_2+O_2 with the MOF structure during static adsorption and kinetic adsorption, we first analyzed the relationship among adsorption selectivity S_{ads} , diffusion selective S_{diff} , and the LCD of $\text{CO}_2/\text{N}_2+\text{O}_2$, as shown in Figure 1. Obviously, most MOFs with large adsorption selectivity and diffusion selectivity have relatively small LCDs. Figure 1a indicates that when the LCD is 2.8–6.5 Å, the adsorption selectivity of $\text{CO}_2/\text{N}_2+\text{O}_2$ decreases, and when the LCD is >15 Å, the adsorption selectivity gradually becomes stable, tending to 5, as depicted by the red line in Figure 1. This is because CO_2 has a strong quadrupole moment, and even in infinitely large pores it is preferentially adsorbed compared to N_2 and O_2 . The trend of gas separation is consistent with the trends of previous reports [49,50]. Figure 1b presents the relationship between S_{diff} and LCD. Similar to S_{ads} , the larger diffusion selectivity ($S_{\text{diff}} > 1$) only occurs in the LCDs ranging from 2.4–5 Å, since the kinetic diameter of CO_2 is less than the kinetic diameters of O_2 and N_2 (the kinetic diameters of CO_2 , O_2 , and N_2 are 3.3, 3.46, and 3.64 Å, respectively). When the LCD of an MOF is small, the CO_2 molecules with smaller diameters diffuse faster, so the diffusion selectivity $S_{\text{diff}(\text{CO}_2/\text{N}_2+\text{O}_2)}$ is larger. As the LCD increases, the diffusion selectivity gradually decreases. When the LCD is >15 Å, the diffusion selectivity tends to be stable and fluctuates around 0.36. Comparison of Figure 1a,b reveals that the adsorption selectivity is generally >1, while it is rare for the diffusion selectivity to be >1. Because CO_2 has a strong quadrupole distance, it has a strong interaction force with MOF molecules, thus hindering the diffusion of CO_2 and resulting in a slower diffusion rate, which may be even smaller than the diffusion rates of N_2 and O_2 .

Figure 1c,d show the relationships of S_{ads} and S_{diff} to the PLD, respectively. Comparing the panels in Figure 1 reveals that the PLD and LCD display the same trend in their relationships to the S_{ads} and S_{diff} of $\text{CO}_2/\text{N}_2+\text{O}_2$. Larger S_{ads} and S_{diff} values appear when the LCD and PLD are small, and S_{ads} and S_{diff} both decrease with increasing PLD or LCD, eventually tending toward stability. Therefore, there is a greater possibility of finding MOFs with simultaneously high S_{ads} and S_{diff} among MOFs with small PLDs and LCDs.

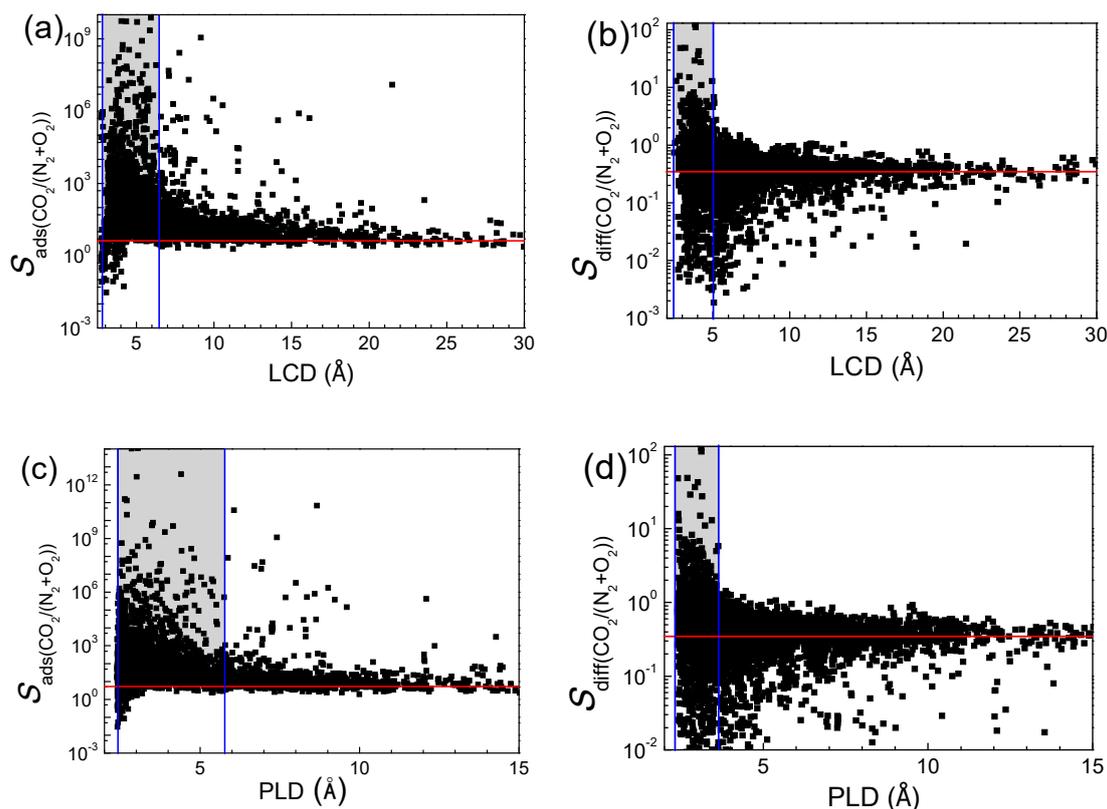


Figure 1. The relationships of (a) $S_{\text{ads}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ and (b) $S_{\text{diff}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ with (c) largest cavity diameter (LCD) and (d) pore-limiting diameter (PLD).

Figure 2a shows that $S_{\text{ads}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ increases monotonically with increasing Q_{st} , indicating that Q_{st} may be the main parameter during the adsorption process. Since the concentration of CO_2 in the atmosphere is low, it is close to the infinite dilution state. Hence, the selectivity is strongly dependent on the isosteric heat of adsorption of CO_2 in the infinite dilution state. The larger $S_{\text{diff}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ in Figure 2b occurs when the VSA is close to zero. As the VSA continues to increase, $S_{\text{diff}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ gradually decreases, and eventually stabilizes. This is because when the VSA is close to zero, the MOF molecule either cannot pass any or only passes a small amount of CO_2 molecules. When the VSA is large, all the gas molecules can pass through the MOF molecule. Therefore, the separation of CO_2 cannot be achieved, i.e., the diffusion selectivity is substantially unchanged. Figure S11b,c indicate the relationship of adsorption selectivity with porosity and VSA, respectively. It can be observed that both of these parameters exert weak influences on adsorption selectivity.

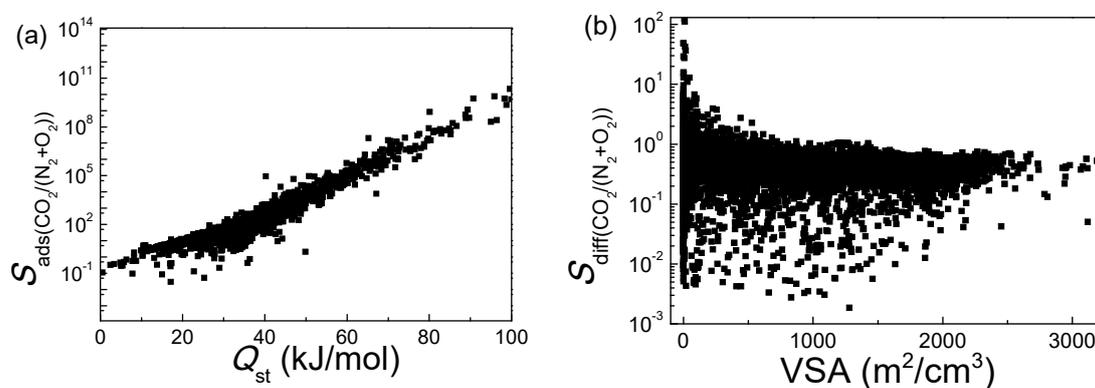


Figure 2. The relationships between (a) $S_{\text{ads}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ and Q_{st} , (b) $S_{\text{diff}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ and volumetric surface area (VSA).

In addition to adsorption and diffusion selectivity, the Henry coefficient of CO₂ reflects the adsorption performance of CO₂ in the infinite dilution state, helping to explain the capture performance of MOFs for air with very low CO₂ concentration. Figure 3a clearly shows the tendency of K_{N_2} to change with enthalpy. When the porosity ϕ is small, the MOF has no space due to the limited pore volume, and only a small amount of N₂ can be adsorbed; therefore, K_{N_2} is small. When ϕ is in the range of 0–0.29, K_{N_2} increases significantly with increasing ϕ . When $\phi > 0.29$, K_{N_2} slows down and gradually stabilizes with increasing ϕ . Figure 3b compares the Henry coefficients of the 3 gases. It can be seen that the trends of the Henry coefficients of N₂ and O₂ are almost the same; however, CO₂ is different. First, in most MOFs, the Henry coefficient values of CO₂ are basically larger than the Henry coefficient values of N₂ and O₂. Second, when LCD > 20 Å, the K_{CO_2} value tends to be level, and eventually stabilizes. The Henry coefficient of CO₂ is still higher than the coefficients of N₂ and O₂, which also leads to MOF selectivity > 1 when the LCD is infinite, as seen in Figure 1a. Finally, it can be observed that only a few MOFs can be identified for which the CO₂ Henry coefficient can be > 10⁻¹ mmol/g/Pa. Observing these MOF structures reveals that most have smaller or open metal sites. The above univariate analysis can only determine the relationship between individual parameters and performance. Q_{st} , PLD and LCD are considered to have dramatic impacts on adsorption selectivity and diffusion selectivity, but their variable influences cannot be analyzed quantitatively. We will further utilize 4 types of ML algorithms to obtain additional information about structure-performance.

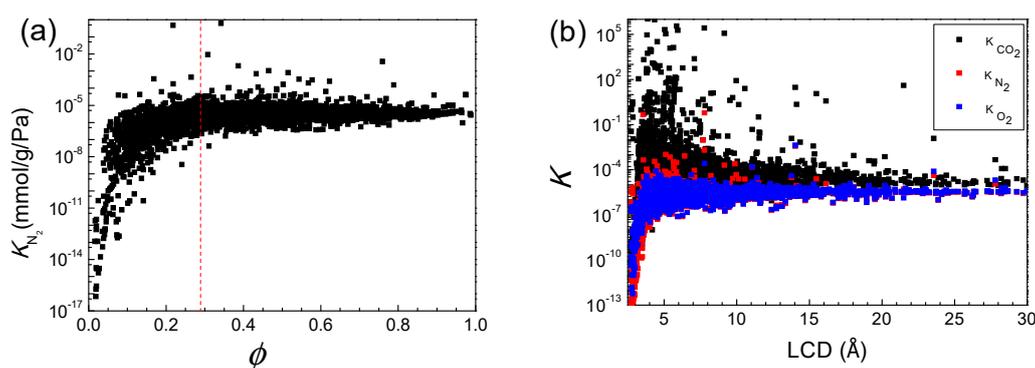


Figure 3. Henry coefficient K_{N_2} versus (a) ϕ and (b) LCD.

3.2. Machine Learning

Currently, machine learning has been used to predict the performance of materials and to filter high-performance materials from large databases [51]. Aiming to discover a better machine prediction method suitable for this system, we individually compared the simulations of the 4 ML algorithms commonly used in big data analysis, i.e., the BPNN, DT, RF, and SVM. Among them, BPNN is a kind of forward signal propagation with error back propagation in which the gradient descent algorithm continuously adjusts the weight and threshold until the error is less than a set threshold. Some parameters of BPNN were set: the training function is Levenberg–Marquardt, the transfer function is a hyperbolic tangent sigmoid transfer function, and the performance evaluation function is the mean square error (MSE). The number of hidden layer neurons was 18, the maximum number of training was 1000, the training required an accuracy of 0.001, and the learning rate was 0.01. DT is a traditional method for data classification and screening. Under the condition that the probability of occurrence takes place in various situations, probability analysis is employed to analyze data with the dendritic model to obtain the expected values. The random forest algorithm is composed of multiple decision trees. The setting parameters of DT were: standard CART (classification and regression tree) used to select the best split predictor at each node. The criteria of splitting and pruning are the MSE function. After optimizing and pruning, the minimum number of branch node observations was 10, the minimum number of leaf node observations was 4, and the maximal number of decision splits was 1. RF uses the method of randomly selecting split attribute sets to construct a decision tree. The parameters for RF were set as: number of trees 200, minimum leaf size 10. The number of variables

randomly selected in the variable subset of the node split in each tree was 2. SVM is an algorithm for binary classification of data through supervised learning, and employs mathematical transformation methods to divide data with a certain centralized structure into rules. We chose the support vector machine regression model of Statistics and Machine Learning Toolbox in MATLAB 2016b to predict, where the kernel function is radial basis function (Gaussian), the kernel scale parameter is set as “auto”, and the loss function is epsilon-insensitive. The box constraint (also called the penalty coefficient, C) was 0.0567, and the half width of epsilon-insensitive band (ϵ) was set as 0.0057. In the radial basis kernel function, $\text{Gamma} = 1/(2\sigma^2)$, where σ is the parameter of the kernel function, which can affect the complexity of the SVM regression algorithm. In our study, the value of gamma was 7.125. The solver of convex quadratic programming is sequential minimal optimization (SMO). Before training and testing, we first processed the data set out-of-order, and then randomly divide it into training and testing sets based on a ratio of 7:3. More detailed descriptions of ML algorithms are listed in the supporting information, and the corresponding diagrams of each algorithm are shown in Figures S2–S5.

We used BPNN, RF, DT, and SVM to predict the adsorption selectivity, and took the logarithm of the adsorption selectivity in order to reduce the differences associated with the varying magnitudes of the data. The 4 ML for predicting the correlation coefficient R value of the adsorption selectivity are listed in Table 1. The results of the testing and training are shown in Figure 4 and Figure S12. The distribution trends of the points in Figure 4a–d are all straight lines inclined upward. The different colors from top to bottom in the figure represent an increase in the number of points, and most of the points are concentrated on the diagonal, indicating that the prediction results are good. Figure 4 reveals that RF has the highest correlation coefficient value (0.982), while the support vector machine algorithm has the lowest (0.886). Thus, the prediction accuracy obtained by the RF algorithm is the highest. Therefore, among the 4 ML algorithms, the structure-performance relationship of RF on adsorption selectivity obtains more information, and the prediction results are the best. RF has good generalization ability and strong model learning ability, and this type of ML is suitable for the system. To verify the accuracy of four ML algorithms, we performed 5 repeated predictions, listed in Table S3. The repeated prediction results do not vary significantly, confirming that the RF algorithm is a suitable model. Because RF introduces two kinds of randomness (sampling randomness and feature randomness), it has strong generalization ability. In previous studies, random forest algorithms also exhibited the best prediction results [52,53]. Whether the overfitting of the model is an important issue. We used combinations of different descriptors and 5 times 5-fold cross-validation to verify the RF model. The results showed that the selected model was not overfitting. During the material screening process, the relative importance of parameters may affect the ultimate screening results. We selected the best RF algorithm to predict the relative importance of each descriptor. The relative importance percentages are shown in Figure 5 and Table S7. We used mean squared error (MSE) to evaluate the relative importance of the 6 descriptors; the greater the percentage of relative importance of the resulting descriptors, the higher the relative contribution of the specific descriptor. According to the results presented in Figure 5, the percentage of Q_{st} is the largest, thus indicating that Q_{st} exerts the greatest impact on the adsorption selectivity. The relative importance of the MOF descriptors to the adsorption selectivity is $Q_{st} > \rho > \text{LCD} > \text{VSA} \geq \phi > \text{PLD}$. The importances of VSA and ϕ are very close, indicating that the effects of two descriptors on adsorption selectivity are roughly equal. From a material science point of view, the larger the ϕ , the larger the VSA. This may be the reason why the effects of these parameters are essentially the same.

Table 1. The 4 ML algorithms for predicting the correlation coefficient R value of the adsorption selectivity.

ML Algorithms	R Value	
	Train	Test
BPNN	0.982	0.979
RF	0.994	0.981
DT	0.985	0.969
SVM	0.915	0.886

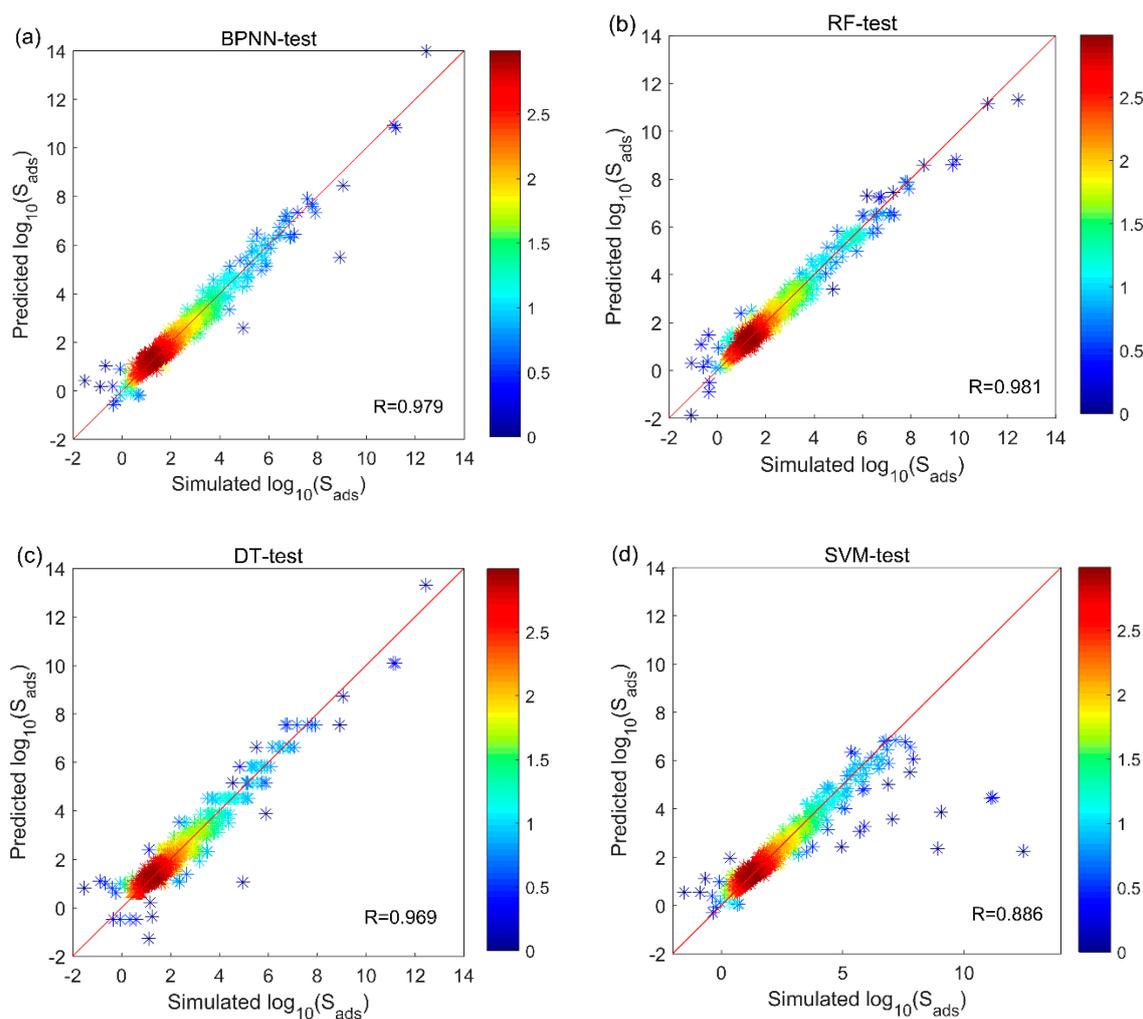


Figure 4. The test results of four machine-learning (ML) algorithms predicted versus simulated $S_{\text{ads}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ using ρ , ϕ , VSA, LCD, PLD and heat of adsorption. (a) back propagation neural network (BPNN), (b) random forest (RF), (c) decision tree (DT), (d) support vector machine (SVM). The color of point represents the number of metal-organic frameworks (MOFs), and the unit of number is a base-10 logarithm of MOF numbers.

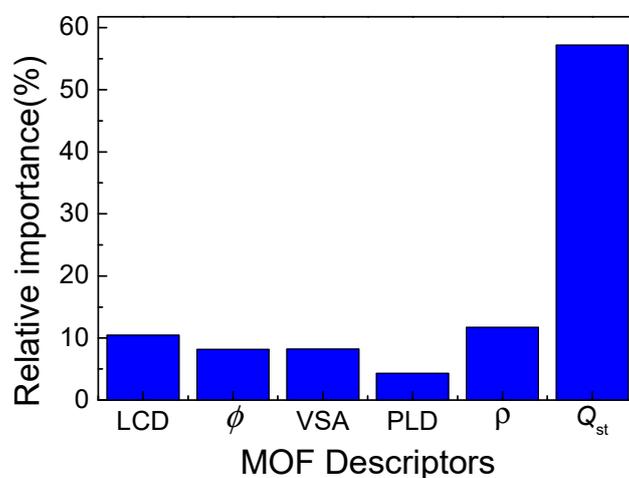


Figure 5. Predicted by the Random Forest, the relative importance of the six descriptors for adsorption selectivity.

4. Best Metal-Organic Frameworks (MOFs)

We selected 5 limiting conditions for $S_{\text{ads}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ and $S_{\text{diff}}(\text{CO}_2/\text{N}_2+\text{O}_2)$, and chose 14 optimal MOFs from the 6013 MOFs, as listed in Table 2 of the 14 materials, HIQPEE exhibited the largest S_{diff} , which was as much as 62.27 Å. NORGOS displayed the largest S_{ads} , which also corresponded to its maximum heat of adsorption. In comparison, the optimal MOF selected by this study is also more selective at higher Q_{st} (4712.33) under the same conditions than that predicted by Wu et al. (433) [30] at $Q_{\text{st}} = 47.8$ kJ/mol and Ravichandar Babarao et al. (500) [54]. It was discovered that diffusion selectivity is generally lower than adsorption selectivity. The diffusion of CO_2 is the key property in determining the performance of MOFs for low concentrations of CO_2 during the kinetic adsorption process. For these 14 MOFs, the LCD, ϕ , and PLD ranges of the six descriptors also corresponded to those in the previous univariate analysis. Especially for PLDs, the optimal range of 2.66–3.64 Å only spans 1 Å, which is also very close to the kinetic diameter of the 3 gases. In such strictly restricted channels, only CO_2 molecules can enter and be adsorbed, greatly increasing the probability of CO_2 being captured at low concentrations. Therefore, the analysis of the optimal MOF revealed that a PLD with a kinetic diameter close to CO_2 is a key condition for good CO_2 diffusion performance, further resulting in the excellent performance of the MOF in capturing CO_2 from the air, and thus providing effective guidance for the design and synthesis of new MOFs.

Table 2. Best computation-ready, experimental metal-organic frameworks (CoRE-MOFs).

No	CSD Code ^a	LCD ^b	ϕ	VSA ^c (m ² /cm ³)	PLD ^d (Å)	P (kg/m ³)	$Q_{\text{st,CO}_2}$ (kJ/mol)	$S_{\text{diffCO}_2/(\text{N}_2+\text{O}_2)}$	$S_{\text{adsCO}_2/(\text{N}_2+\text{O}_2)}$
1	REYCEF	3.75	0.08	0	2.83	1646.40	26.02	8.21	6.25
2	JAHNEM	3.50	0.05	0	2.66	1714.62	27.01	6.18	6.77
3	HOJLEY	3.63	0.14	0	2.83	1410.99	32.65	5.00	7.45
4	KASPOL	4.04	0.19	16.26	2.93	1665.95	34.05	36.55	7.68
5	XUNJOG	3.25	0.07	0	2.67	1737.75	26.41	5.31	11.12
6	HIQPII	3.87	0.15	9.46	3.12	1472.40	35.20	55.67	11.24
7	YUBFUX	4.58	0.16	98.56	3.64	1786.84	30.52	5.79	11.51
8	HIQPEE	3.84	0.15	7.68	3.12	1440.14	35.31	62.27	12.39
9	FEJKEM	3.46	0.09	0.33	3.09	2132.02	30.00	15.01	15.72
10	FALQIU	5.08	0.12	82.58	3.14	1977.64	34.40	5.38	22.97
11	FALQOA	6.06	0.12	83.34	3.07	2004.41	35.03	6.76	24.10
12	TOXNAX	3.80	0.28	4.42	2.83	1346.52	42.65	5.02	25.15
13	OFIWIK	4.20	0.05	5.99	3.14	1866.86	39.64	27.42	25.66
14	NORGOS	4.95	0.13	45.40	3.43	1728.90	51.67	12.92	4712.33

^a CSD Code is the code of MOFs in the Cambridge Structure Database; ^b LCD: largest cavity diameter; ^c VSA: volumetric surface area; ^d PLD: pore-limiting diameter.

5. Conclusions

Firstly, we simulated the adsorption and diffusion properties of CO_2 , N_2 , and O_2 in 6013 CoRE-MOFs using high-throughput MC + MD. Then, we investigated the correlation among adsorption selectivity and diffusion selectivity for CO_2 and MOF descriptors by the univariate analysis. The Q_{st} and PLD were considered to be the most important for $S_{\text{ads}}(\text{CO}_2/\text{N}_2+\text{O}_2)$ and $S_{\text{diff}}(\text{CO}_2/\text{N}_2+\text{O}_2)$, respectively. In conjunction with multivariate analysis, a comparison of 4 ML algorithms revealed that the RF had the best prediction results for adsorption selectivity, with an R value of 0.982. This indicated that the RF method was the most suitable for the predictions of the capture of low CO_2 concentrations in MOF. The relative importance analysis of the RF algorithm quantitatively indicated that the relative importance of the MOF descriptors on adsorption selectivity is $Q_{\text{st}} > \rho > \text{LCD} > \text{VSA} \geq \phi > \text{PLD}$. It was also confirmed that Q_{st} is the most important parameter, while the VSA and ϕ are relatively less important. Through this high-throughput screening, 14 types of MOFs with optimal adsorption selectivity and diffusion selectivity were obtained. After comparison, it was found that their adsorption selectivity was generally higher than their diffusion selectivity. The diffusion separation performance of CO_2 is the key property in determining the performance of MOFs on low concentrations of CO_2 during

the kinetic adsorption process. This study provides experimental guidance for the determination of MOFs that effectively capture CO₂ from the air, and indicates that advanced ML algorithms can accelerate the research and development of new materials.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2076-3417/10/2/569/s1>: Figure S1: Models of N₂, O₂ and CO₂, Figure S2: Back propagation neural network, Figure S3: Decision tree, Figure S4: Random forest, Figure S5: Support vector machine, Figure S6: Diffusion coefficient D_i versus (a) LCD, (b–d) Q_{st} (D_i : the diffusion coefficient of different gases, (a,b) CO₂, (c) N₂, (d) O₂), Figure S7: Henry coefficient K_{CO_2} versus (a) ϕ , (b) PLD, (c) Q_{st} and (d) VSA, Figure S8: Henry coefficient K_{N_2} versus (a) ϕ , (b) PLD, (c) Q_{st} and (d) VSA, Figure S9: Henry coefficient K_{O_2} versus (a) ϕ , (b) PLD, (c) Q_{st} and (d) VSA, Figure S10: Diffusion selectivity $S_{diff(CO_2/(N_2+O_2))}$ versus (a) LCD, (b) ϕ , (c) PLD and (d) VSA, Figure S11: Adsorption selectivity $S_{ads(CO_2/N_2+O_2)}$ versus (a) LCD, (b) ϕ , (c) PLD and (d) VSA, Figure S11: The training results of four machine learning predicted versus MC simulated $\log_{10}(S_{ads(CO_2/N_2+O_2)})$ using density, void fraction, volumetric surface area, density, LCD, PLD and heat of adsorption. (a) BPNN, (b) RF, (c) DT and (d) SVM. The color of point represents the number of MOF, and the unit of number is a base-10 logarithm of MOF numbers, Table S1: Lennard-Jones parameters of MOFs, Table S2: Lennard-Jones parameters and charges of adsorbates, Table S3: The training and testing R values of adsorption selectivity using repeat 5 time-four ML. Table S4 Prediction using RF models with different descriptor combinations. Table S5 Prediction using repeat 5 times-RF models with different descriptor combinations. Table S6 The results of predicted RF with k times k -fold cross validation. Table S7: Predicted by the RF the relative importance of the six descriptors for adsorption selectivity. Formula S1: $S_{ads(CO_2/N_2+O_2)}$ indicates the adsorption selectivity of CO₂/N₂+O₂; K_i represents the Henry coefficient of component i (CO₂, N₂ and O₂), Formula S2: $S_{diff(CO_2/(N_2+O_2))}$ represents the diffusion selectivity of CO₂/N₂+O₂; D_i represents the diffusion coefficient of component i .

Author Contributions: X.D., Z.S., H.L. and Z.Q. conceived the idea. Z.Q. calculated all the materials' structural parameters and obtained valid data about the structure descriptors and performance. X.D., W.Y. and Z.S. analyzed the relationship between structure descriptors and performance. X.D. and S.L. used univariate analysis to obtain the influence law of affecting CO₂ adsorption and diffusion in air and Z.S. used ML algorithms to predict the MOF performance. X.D. and Z.S. wrote the original draft. H.L. and Z.Q. wrote the manuscript with contributions from all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Nos. 21978058, 21676094, and 21576058).

Acknowledgments: We gratefully thank the National Natural Science Foundation of China (Nos. 21978058, 21676094, and 21576058) for financial support.

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

References

1. Canadell, J.G.; Le Quere, C.; Raupach, M.R.; Field, C.B.; Buitenhuis, E.T.; Ciais, P.; Conway, T.J.; Gillett, N.P.; Houghton, R.A.; Marland, G. Contributions to accelerating atmospheric CO₂ growth from economic activity, carbon intensity, and efficiency of natural sinks. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 18866–18870. [[CrossRef](#)] [[PubMed](#)]
2. Zhang, C.; Zeng, G.; Huang, D.; Lai, C.; Chen, M.; Cheng, M.; Tang, W.; Tang, L.; Dong, H.; Huang, B.; et al. Biochar for environmental management: Mitigating greenhouse gas emissions, contaminant treatment, and potential negative impacts. *Chem. Eng. J.* **2019**, *373*, 902–922. [[CrossRef](#)]
3. Fan, R.; Chen, C.-L.; Lin, J.-Y.; Tzeng, J.-H.; Huang, C.-P.; Dong, C.; Huang, C.P. Adsorption characteristics of ammonium ion onto hydrous biochars in dilute aqueous solutions. *Bioresour. Technol.* **2019**, *272*, 465–472. [[CrossRef](#)] [[PubMed](#)]
4. Fang, G.; Liu, C.; Wang, Y.; Dionysiou, D.D.; Zhou, D. Photogeneration of reactive oxygen species from biochar suspension for diethyl phthalate degradation. *Appl. Catal. B Environ.* **2017**, *214*, 34–45. [[CrossRef](#)]
5. Available online: <https://carbonengineering.com/> (accessed on 10 July 2019).
6. Nibleus, K.; Lundin, R. Climate Change and Mitigation. *Ambio* **2010**, *39*, 11–17. [[CrossRef](#)]
7. Boyd, P.G.; Chidambaram, A.; García-Díez, E.; Ireland, C.P.; Daff, T.D.; Bounds, R.; Gładysiak, A.; Schouwink, P.; Moosavi, S.M.; Maroto-Valer, M.M.; et al. Data-driven design of metal–organic frameworks for wet flue gas CO₂ capture. *Nature* **2019**, *576*, 253–256. [[CrossRef](#)]
8. Faig, R.W.; Popp, T.M.O.; Fracaroli, A.M.; Kapustin, E.A.; Kalmutzki, M.J.; Altamimi, R.M.; Fathieh, F.; Reimer, J.A.; Yaghi, O.M. The Chemistry of CO₂ Capture in an Amine-Functionalized Metal-Organic Framework under Dry and Humid Conditions. *J. Am. Chem. Soc.* **2017**, *139*, 12125–12128. [[CrossRef](#)]

9. Haszeldine, R.S. Carbon Capture and Storage: How Green Can Black Be? *Science* **2009**, *325*, 1647–1652. [[CrossRef](#)]
10. McDonald, T.M.; Mason, J.A.; Kong, X.; Bloch, E.D.; Gygi, D.; Dani, A.; Crocellà, V.; Giordanino, F.; Odoh, S.O.; Drisdell, W.S.; et al. Cooperative insertion of CO₂ in diamine-appended metal-organic frameworks. *Nature* **2015**, *519*, 303–308. [[CrossRef](#)]
11. Liu, J.; Wei, Y.; Zhao, Y. Trace Carbon Dioxide Capture by Metal-Organic Frameworks. *ACS Sustain. Chem. Eng.* **2019**, *7*, 82–93. [[CrossRef](#)]
12. Zhao, R.; Liu, L.; Zhao, L.; Deng, S.; Li, S.; Zhang, Y.; Li, H. Thermodynamic exploration of temperature vacuum swing adsorption for direct air capture of carbon dioxide in buildings. *Energy Convers. Manag.* **2019**, *183*, 418–426. [[CrossRef](#)]
13. Batten, S.R.; Champness, N.R.; Chen, X.-M.; Garcia-Martinez, J.; Kitagawa, S.; Ohrstrom, L.; O’Keeffe, M.; Suh, M.P.; Reedijk, J. Terminology of metal-organic frameworks and coordination polymers (IUPAC Recommendations 2013). *Pure Appl. Chem.* **2013**, *85*, 1715–1724. [[CrossRef](#)]
14. Murray, L.J.; Dinca, M.; Long, J.R. Hydrogen storage in metal-organic frameworks. *Chem. Soc. Rev.* **2009**, *38*, 1294–1314. [[CrossRef](#)] [[PubMed](#)]
15. Sculley, J.; Yuan, D.; Zhou, H.-C. The current status of hydrogen storage in metal-organic frameworks—updated. *Energy Environ. Sci.* **2011**, *4*, 2721–2735. [[CrossRef](#)]
16. Li, J.-R.; Kuppler, R.J.; Zhou, H.-C. Selective gas adsorption and separation in metal-organic frameworks. *Chem. Soc. Rev.* **2009**, *38*, 1477–1504. [[CrossRef](#)] [[PubMed](#)]
17. Verma, S.; Mishra, A.K.; Kumar, J. The Many Facets of Adenine: Coordination, Crystal Patterns, and Catalysis. *Acc. Chem. Res.* **2010**, *43*, 79–91. [[CrossRef](#)]
18. Li, J.-R.; Sculley, J.; Zhou, H.-C. Metal-Organic Frameworks for Separations. *Chem. Rev.* **2012**, *112*, 869–932. [[CrossRef](#)]
19. Bae, Y.-S.; Snurr, R.Q. Development and Evaluation of Porous Materials for Carbon Dioxide Separation and Capture. *Angew. Chem. Int. Ed.* **2011**, *50*, 11586–11596. [[CrossRef](#)]
20. Wu, X.-J.; Zhao, P.; Fang, J.-M.; Wang, J.; Liu, B.-S.; Cai, W.-Q. Simulation on the Hydrogen Storage Properties of New Doping Porous Aromatic Frameworks. *Acta Phys. Chim. Sin.* **2014**, *30*, 2043–2054. [[CrossRef](#)]
21. Wu, P.; He, C.; Wang, J.; Peng, X.; Li, X.; An, Y.; Duan, C. Photoactive Chiral Metal-Organic Frameworks for Light-Driven Asymmetric α -Alkylation of Aldehydes. *J. Am. Chem. Soc.* **2012**, *134*, 14991–14999. [[CrossRef](#)]
22. Farrusseng, D.; Aguado, S.; Pinel, C. Metal-Organic Frameworks: Opportunities for Catalysis. *Angew. Chem. Int. Ed.* **2009**, *48*, 7502–7513. [[CrossRef](#)] [[PubMed](#)]
23. Ma, L.; Abney, C.; Lin, W. Enantioselective catalysis with homochiral metal-organic frameworks. *Chem. Soc. Rev.* **2009**, *38*, 1248–1256. [[CrossRef](#)] [[PubMed](#)]
24. Lee, J.; Farha, O.K.; Roberts, J.; Scheidt, K.A.; Nguyen, S.T.; Hupp, J.T. Metal-organic framework materials as catalysts. *Chem. Soc. Rev.* **2009**, *38*, 1450–1459. [[CrossRef](#)] [[PubMed](#)]
25. Farha, O.K.; Shultz, A.M.; Sarjeant, A.A.; Nguyen, S.T.; Hupp, J.T. Active-Site-Accessible, Porphyrinic Metal-Organic Framework Materials. *J. Am. Chem. Soc.* **2011**, *133*, 5652–5655. [[CrossRef](#)] [[PubMed](#)]
26. Della Rocca, J.; Liu, D.; Lin, W. Nanoscale Metal-Organic Frameworks for Biomedical Imaging and Drug Delivery. *Acc. Chem. Res.* **2011**, *44*, 957–968. [[CrossRef](#)]
27. Bernini, M.C.; Fairen-Jimenez, D.; Pasinetti, M.; Ramirez-Pastor, A.J.; Snurr, R.Q. Screening of bio-compatible metal-organic frameworks as potential drug carriers using Monte Carlo simulations. *J. Mater. Chem. B* **2014**, *2*, 766–774. [[CrossRef](#)]
28. Peng, Y.-W.; Wu, R.-J.; Liu, M.; Yao, S.; Geng, A.-F.; Zhang, Z.-M. Nitrogen Coordination to Dramatically Enhance the Stability of In-MOF for Selectively Capturing CO₂ from a CO₂/N₂ Mixture. *Cryst. Growth Des.* **2019**, *19*, 1322–1328. [[CrossRef](#)]
29. Shekhah, O.; Belmabkhout, Y.; Chen, Z.; Guillerm, V.; Cairns, A.; Adil, K.; Eddaoudi, M. Made-to-order metal-organic frameworks for trace carbon dioxide removal and air capture. *Nat. Commun.* **2014**, *5*. [[CrossRef](#)]
30. Jain, A.; Shyue Ping, O.; Hautier, G.; Chen, W.; Richards, W.D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **2013**, *1*. [[CrossRef](#)]

31. Furukawa, H.; Cordova, K.E.; O’Keeffe, M.; Yaghi, O.M. The Chemistry and Applications of Metal-Organic Frameworks. *Science* **2013**, *341*, 1230444. [[CrossRef](#)]
32. 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](#)] [[PubMed](#)]
33. Lin, L.-C.; Berger, A.H.; Martin, R.L.; Kim, J.; Swisher, J.A.; Jariwala, K.; Rycroft, C.H.; Bhowan, A.S.; Deem, M.W.; Haranczyk, M.; et al. In silico screening of carbon-capture materials. *Nat. Mater.* **2012**, *11*, 633–641. [[CrossRef](#)] [[PubMed](#)]
34. Wu, D.; Yang, Q.; Zhong, C.; Liu, D.; Huang, H.; Zhang, W.; Maurin, G. Revealing the Structure-Property Relationships of Metal-Organic Frameworks for CO₂ Capture from Flue Gas. *Langmuir* **2012**, *28*, 12094–12099. [[CrossRef](#)] [[PubMed](#)]
35. Fernandez, M.; Boyd, P.G.; Daff, T.D.; Aghaji, M.Z.; Woo, T.K. Rapid and Accurate Machine Learning Recognition of High Performing Metal Organic Frameworks for CO₂ Capture. *J. Phys. Chem. Lett.* **2014**, *5*, 3056–3060. [[CrossRef](#)] [[PubMed](#)]
36. Available online: <https://github.com/gregchung/gregchung.github.io/blob/master/CoRE-MOFs/structure-doi-CoRE-MOFsV2.0.csv> (accessed on 15 January 2019).
37. 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](#)]
38. Willems, T.F.; Rycroft, C.; Kazi, M.; Meza, J.C.; Haranczyk, M. Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous Mesoporous Mater.* **2012**, *149*, 134–141. [[CrossRef](#)]
39. Dubbeldam, D.; Calero, S.; Ellis, D.E.; Snurr, R.Q. RASPA: Molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Mol. Simul.* **2016**, *42*, 81–101. [[CrossRef](#)]
40. 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](#)]
41. Potoff, J.J.; Siepmann, J.I. Vapor-liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen. *AIChE J.* **2001**, *47*, 1676–1682. [[CrossRef](#)]
42. Kadantsev, E.S.; Boyd, P.G.; Daff, T.D.; Woo, T.K. Fast and Accurate Electrostatics in Metal Organic Frameworks with a Robust Charge Equilibration Parameterization for High-Throughput Virtual Screening of Gas Adsorption. *J. Phys. Chem. Lett.* **2013**, *4*, 3056–3061. [[CrossRef](#)]
43. 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](#)]
44. 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](#)]
45. Bian, L.; Li, W.; Wei, Z.; Liu, X.; Li, S. Formaldehyde Adsorption Performance of Selected Metal-Organic Frameworks from High-throughput Computational Screening. *Acta Chim. Sin.* **2018**, *76*, 303–310. [[CrossRef](#)]
46. Rappe, A.K.; Casewit, C.J.; Colwell, K.S.; Goddard, W.A.; 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](#)]
47. 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](#)]
48. Babarao, R.; Jiang, J. Diffusion and separation of CO₂ and CH₄ in silicalite, C₁₆₈ schwarzite, and IRMOF-1: A comparative study from molecular dynamics simulation. *Langmuir* **2008**, *24*, 5474–5484. [[CrossRef](#)]
49. Qiao, Z.; Zhang, K.; Jiang, J. In silico screening of 4764 computation-ready, experimental metal-organic frameworks for CO₂ separation. *J. Mater. Chem. A* **2016**, *4*, 2105–2114. [[CrossRef](#)]
50. 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. *Energy Environ. Sci.* **2012**, *5*, 9849–9856. [[CrossRef](#)]
51. Takahashi, K.; Tanaka, Y. Materials informatics: A journey towards material design and synthesis. *Dalton Trans.* **2016**, *45*, 10497–10499. [[CrossRef](#)]
52. Pardakhti, M.; Moharreri, E.; Wanik, D.; Suib, S.L.; Srivastava, R. Machine Learning Using Combined Structural and Chemical Descriptors for Prediction of Methane Adsorption Performance of Metal Organic Frameworks (MOFs). *ACS Comb. Sci.* **2017**, *19*, 640–645. [[CrossRef](#)]

53. 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](#)]
54. Babarao, R.; Jiang, J. Unprecedentedly High Selective Adsorption of Gas Mixtures in rho Zeolite-like Metal-Organic Framework: A Molecular Simulation Study. *J. Am. Chem. Soc.* **2009**, *131*, 11417–11425. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).