

研究论文

DOI: 10.11949/0438-1157.

基于数据驱动和材料指纹的低温储氢 MOF 筛选

蔡嘉纯, 黄河, 田济铜, 龚凌涛, 黄晓珊, 林颖潼, 李树华, 梁红, 施俊越¹, 吴玉芳¹, 乔智威¹¹ 广州大学化学化工学院, 广东 广州 510006

摘要: 为利用大数据精准分析与开发高性能低温储氢吸附材料, 本研究利用 MACCS 指纹将数万个已合成金属-有机框架 (Metal-Organic Framework, MOF) 数据化, 基于 SHAP 技术和 5 种机器学习算法构建预测模型, 结合高通量分子模拟结果, 揭示了 MOF 材料在低温高压环境 (159K、30bar) 存储氢气的构效关系, 其中分类提升算法 (Categorical Boosting, CatBoost) 的预测效果最好 (R^2 可达 0.951)。MOF 材料指纹对工作容量 (ΔW) 贡献度的研究表明, 第 99、62、45、65 号指纹是提升 MOF 低温储氢吸附能力的关键子结构, 其中第 65 号指纹代表芳香碳氮杂环 (如咪唑、吡啶等), 证实当 MOF 性能已达到 TOP10% 时, 通过引入芳香碳氮杂环, 可在低温下局部形成“类氢键”, 再次突破 MOF 储氢性能 (可达 TOP5%)。

关键词: 氢; 吸附; 分子模拟; 机器学习; 分子指纹; 金属有机框架

中图分类号: TP181; TG174.42

文献标志码: A

文章编号: 0438-1157 (XXXX) XX-0001-12

Screening of low-temperature hydrogen storage MOFs via data-driven approaches and material fingerprints

CAI Jiachun, HUANG He, TIAN Jitong, GONG Lingtao, HUANG Xiaoshan, LIN Yingtong, LI Shuhua, LIANG Hong, SHI Junyue¹, WU Yufang¹, QIAO Zhiwei¹¹ College of Chemistry and Chemical Engineering, Guangzhou University, Guangzhou 510006, Guangdong, China

Abstract: This study employs a data-driven strategy to enable precise analysis and the development of high-performance cryogenic hydrogen storage adsorbent materials. In this work, thousands of Computation-Ready Experimental MOF (CoRE-MOF) were digitized using molecular fingerprints. Predictive models were constructed based on SHAP technology and five machine learning algorithms, and integrated with high-throughput molecular simulation results to elucidate the structure-property relationship of MOF materials for hydrogen storage under cryogenic high-pressure conditions (159K and 30bar). Among the models, the categorical boosting algorithm (CatBoost) achieved the best predictive performance, with R^2 of 0.951. The results indicate that three descriptors—adsorption heat (Q_{st}), density (ρ), and porosity (φ)—are critical for enhancing material adsorption performance. Specifically, when $Q_{st} > 15 \text{ kJ} \cdot \text{mol}^{-1}$, $\rho < 1000 \text{ kg} \cdot \text{m}^{-3}$, and $\varphi > 0.7$, MOFs exhibit a high level of adsorption performance.

收稿日期: 2026-04-10 修回日期: 2026-05-06

通信作者: 乔智威 (1986—), 男, 博士, 教授, zqiao@gzhu.edu.cn; 吴玉芳 (1992—), 女, 博士, 副教授, yufang.wu@gzhu.edu.cn; 施俊越 (1999—), 男, 在读研究生, 112405096@gzhu.edu.cn

第一作者: 蔡嘉纯 (2004—), 女, 在读本科生, 32305200012@e.gzhu.edu.cn

基金项目: 国家自然科学基金项目 (22478085, 22308069, 21978058); 广东省自然科学基金项目 (2023A1515240076, 2022A1515011446); 新型反应器与绿色化学工艺湖北省重点实验室开放基金 (NRG202407); 广州大学校级大学生创新创业项目 (XJ202511078241)

引用本文: 蔡嘉纯, 黄河, 田济铜, 龚凌涛, 黄晓珊, 林颖潼, 李树华, 梁红, 施俊越, 吴玉芳, 乔智威. 基于数据驱动和材料指纹的低温储氢 MOF 筛选[J]. 化工学报, XXXX, XX(XX): 1–12

Citation: CAI Jiachun, HUANG He, TIAN Jitong, GONG Lingtao, HUANG Xiaoshan, LIN Yingtong, LI Shuhua, LIANG Hong, SHI Junyue, WU Yufang, QIAO Zhiwei. Screening of low-temperature hydrogen storage MOFs via data-driven approaches and material fingerprints[J]. CIESC Journal, XXXX, XX(XX): 1–12

Analysis of the variation in RI across different datasets reveals that as material performance reaches the top 20% or top 10%, the influence of descriptors such as Q_{st} on MOF adsorption performance gradually diminishes, while the contribution of molecular fingerprints to ΔW becomes increasingly significant. Therefore, this study employs SHAP analysis to interpret the model constructed using the CatBoost algorithm, quantifying the contribution of each molecular fingerprint to cryogenic hydrogen storage performance (ΔW). The final results identify fingerprints 99, 62, 45, and 65 as key substructures for enhancing hydrogen storage capacity in MOFs. Among these, fingerprint 65 corresponds to aromatic nitrogen-containing heterocycles (such as imidazole, pyridine), confirming that when MOF performance reaches the top 10%, the introduction of such aromatic nitrogen heterocycles facilitates the formation of local hydrogen bond-like interactions under cryogenic conditions, further boosting hydrogen storage performance to the top 5%. Through this data-driven strategy, this study precisely identifies the key substructures governing the cryogenic hydrogen storage performance of MOFs, laying a solid foundation for the design of high-performance cryogenic hydrogen storage materials.

Keywords: hydrogen; adsorption; molecular simulation; machine learning; molecular fingerprint; metal-organic frameworks

引 言

氢能具有零碳排、高效、可再生等突出优势^[1],其在缓解气候变化中的主要作用是替代难以电气化的化石燃料^[2],被认为是很有前景的化石能源替代品,可以帮助全球能源体系从以化石燃料为主的传统体系逐步向高效、可再生的低碳能源体系转型^[3]。目前海上风电制氢技术是提高风能利用率的有效途径,既缓解了风电消纳的问题,又实现了氢能产业链的无碳化^[4]。然而,如何利用管道或船舶将氢气输送到用氢地,需要先进、高效且安全环保^[5]的氢气储存技术。

商业化的氢气储存技术可分为气态氢、液态氢和固态氢三种储存方式,其中气态氢储存是将氢气压缩存储至耐高压的容器中,这种方法具有较为成熟的工艺且充放氢气效率高,但对储罐的材质要求较高且存在氢气泄漏和容器爆炸的风险。液态氢储存则是在低温高压条件下将氢气液化,液态氢的密度比气态氢高且运输更为简单,但从气态氢转化为液态氢所需的能耗较高、价格昂贵。与前两者相比,固态氢储存因其较高的氢气储存能力、适宜的工作温度和工作压力以及较低的安全风险,具有实现“双碳”目标的巨大潜力。因此,发现、开发和优化高性能固态氢储存材料,是未来氢气存储和氢能高效利用的关键所在^[6-8]。

固态氢储存方法可根据吸附力分为物理吸附和化学吸附,物理吸附依赖相对较弱的范德华力进行结合,而化学吸附则依赖更强的化学键结合。常

用的化学吸附材料主要包括金属氢化物、配位氢化物、液态有机氢化物等^[9],而物理吸附材料主要包括金属有机框架(Metal-Organic Frameworks, MOFs)^[10]、各种沸石^[11]和活性炭^[12]。其中MOFs是一类崭新的有机-无机杂化高分子材料,由金属离子或金属团簇与有机配体通过配位键连接构成^[13]。与沸石、活性炭等吸附材料相比,MOFs因其极高的比表面积和孔隙率以及结构和性质的可调性,在气体存储^[14-16]、催化^[17]以及分离^[18-20]等领域展现出巨大的应用前景。特别是在气体储存应用方面,这些特性使得MOFs能够高效吸附氢分子,具有卓越的氢气储存效率。同时,在运输氢气的过程中,采用高效吸附材料可以显著降低相关的安全风险。

LIU等人^[21]研究发现通过温和退火Mg-MOF-74制备一种高度缺陷的骨架材料de-MgMOF,通过利用Pt溢出催化剂,该材料在160℃、81bar下具备快速的氢气吸收/释放能力,展现出2.6 wt%的卓越氢气储存能力。CHEN等人^[22]利用温和的刻蚀方法在缺陷型MOF-801的锆(Zr)周围引入过量电荷,形成具有优异吸附热的特定缺陷位点,诱导氢气极化,从而提升储氢容量。在77 K和80 bar条件下,其总氢气吸附量从5.02 wt%(MOF-801)增加至8.58 wt%(Cl@MOF-801)。ZHAO等人^[23]的研究工作表明,通过Ce修饰和MoS₂掺杂的协同策略,可显著提升MOF-808的储氢性能。其中,以Ce:Zr=2:4比例合成的Ce-MOF-808为基础,进一步复合MoS₂所得材料表现出最优的综合性能。这些研究不仅加深了对MOFs材料设计的理解,更为开发高效的氢气存储

系统指明了关键的优化方向与实用路径。

然而,随着 MOFs 研究领域的迅速发展,目前已经实际合成的 MOFs 加上计算机假设模拟得到的 MOFs,其总数已经达到了几十万种^[24]。在如此庞大的 MOFs 数据库里面,仅通过实验去获得高性能的 MOFs 储氢性能材料无异于是大海捞针,这个过程不仅会耗费大量的人力物力财力,其得到的材料也未必能比现有的材料性能更加优异。在计算模拟中,利用机器学习(Machine Learning, ML)和高通量计算筛选(High-Throughput Computational Screening, HTCS) MOFs 往往被认为是更有效的方法^[25-27]。JANG 等人^[28]开发了一种名为 PU-CGCNN 的合成可行性预测机器学习模型,该模型采用正例-无标签学习(PU learning)和晶体图卷积神经网络(CGCNN)技术。Wan 等人^[29]利用 PU-CGCNN 预测得到具有氢气储存潜力且可能合成的 MOFs,并采用机器学习预测的可合成微函数结构进行高通量筛选,随后进行大典范蒙特卡洛(GCMC)和密度泛函理论(DFT)计算,再进行材料合成。最后成功合成一种新型钒基 MOFs,该 MOFs 在低温氢气储存方面表现出卓越的性能,值得注意的是,在 77K 和 150bar 的条件下,其总重量储氢密度和体积储氢密度分别高达 9.0 wt% 和 50.0 g·L⁻¹。WANG 等人^[30]结合 GCMC 模拟和 6 种基于树模型的 ML 算法对 CORE MOF 数据库中的材料进行了储氢性能筛选,以识别那些有潜力超越高压压缩储氢的 MOFs。这项研究结果表明大量模型和实验结果在低温条件下仍然可用,且确定了 362 种在备用电源应用中性能可优于 350bar 压缩氢气储能的 MOFs。

这种 ML 辅助的材料设计为开发高性能的储氢材料提供了一种高效且前景广阔的方法^[31,32],然而针对几乎无限的 MOFs 数据库,在高通量计算过程中,大量无效的计算被运行,效率相对低下。因此,针对 MOFs 材料的储氢吸附性能研究一种合理快速的材料设计方法,是目前该领域的重点研究方向。在这项工作中,通过 HTCS 和分子模拟技术,探究数万方 MOFs 在低温条件下的储氢吸附性能。结合单变量分析,揭示 MOFs 的 7 个特征描述符(体积比表面积(VSA, m²·cm⁻³)、孔隙率(φ)、最大空腔直径(LCD, Å)、孔径限制直径(PLD, Å)、密度(ρ , kg·m⁻³)、吸附热(Q_{st} , kJ·mol⁻¹)、亨利系数(K , mol·kg⁻¹·Pa⁻¹)与其吸附性能之间的关系。随后采用 5 种 ML 算法对 MOFs 的吸附性能进行预测分析,选择预测

效果最好的 ML 算法构建吸附模型并结合 SHAP 技术分析,进一步探究特征描述符和分子指纹(Molecular Fingerprint, MF)对不同数据集中的 MOFs 吸附性能的贡献度。最后探究 MF 对高吸附性能 MOFs 数据集的影响效果,并筛选出前 14 个最优指纹,为开发高性能的低温储氢吸附材料提出新的设计思路。

1 模型与方法

1.1 模型

为设计出高性能的低温储氢吸附材料,基于 CHUANG 等人^[33]开发的 Computation-Ready Experimental MOF(CoRE-MOF) 2019 数据库,本研究对 14142 种 MOFs 进行高通量计算筛选。值得注意的是,该数据库在原始实验结构的基础上,经过了一系列标准化的预处理,包括去除自由溶剂分子和配位溶剂分子、统一晶体对称性、保留具有有效孔隙的结构,并剔除了无序及重复的样本,最终得到适用于机器学习建模的高质量 MOF 结构数据集。所以 MOFs 均采用 7 个特征描述符进行表征,包括 VSA、 φ 、LCD、PLD、 ρ 、 Q_{st} 和 K 。其中,PLD、LCD 和 ρ 使用开源程序 Zeo++ 计算得到,而 φ 、VSA 则是分别以氦气(He: 动力学直径=2.58 Å)和氮气(N₂: 动力学直径=3.64 Å)作为探针通过 RASPA 软件计算得到。 Q_{st} 、 K 作为能量描述符,采用 NTV-MC 方案在 RASPA 软件中进行模拟计算得到^[34]。为了模拟该体系在不同化学势下的吸附行为,采用 Lennard-Jones(LJ)参数和静电势来描述 MOFs 与吸附质之间的相互作用力:

$$u_{LJ} + elec(r) = \sum 4\epsilon_{ij} \left[\left(\frac{\delta_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\delta_{ij}}{r_{ij}} \right)^6 \right] + \sum \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (1)$$

在公式(1)中, q_i 和 q_j 分别是原子*i*和*j*的原子电荷, $\epsilon_0 = 8.8542 \times 10^{-12} \text{C}^2 \cdot \text{N}^{-1} \cdot \text{m}^{-2}$,表示真空介电常数, δ_{ij} 表示原子相互碰撞时的最小距离, r_{ij} 是原子对的距离, ϵ_{ij} 表示能量最低点到能量为零的差。所有 MOFs 骨架原子的 LJ 势能参数均来自通用力场(UFF)和 Dreiding 力场(详见表 S1)^[35]。在分子模拟研究中,采用三位点势能模型描述 H₂ 分子:两个 H 原子上分别带部分电荷($q_H = 0.468e$),质心点带部分电荷($q_{com} = -0.936e$)^[36]。H₂ 分子采用刚性结构处理,两个 H 原子间键长为 0.074 nm^[37]。另外,为了进一步验证模拟力场对 H₂ 分子在 MOF 中吸附行为为预

测的准确性,我们选取了三种具有代表性的 MOF 材料,它们在孔道结构及化学环境方面均存在一定差异。在这项工作中,我们模拟了三种材料对 H_2 的吸附等温线,并将模拟结果与文献中的实验数据进行系统对比^[38, 39]。如图 S1 所示,模拟得到的吸附等温线与实验数据基本吻合,验证了该力场模型的准确性和可靠性。

1.2 分子模拟

在这项工作中,使用巨正则蒙特卡洛(GCMC)方法,分别模拟在 159K、30bar 条件下 MOFs 吸附氢气的过程以及在 298K、1bar 条件下 MOFs 解吸氢气的过程。本次研究采用 RASPA 软件进行 GCMC 模拟,其中每个 GCMC 模拟计算分别对每个 MOFs 独立运行,且整个模拟过程都在恒定的化学势、体积和温度条件下进行,MOFs 骨架与吸附分子之间的交叉相互作用通过 Lorentz-Berthelot 组合规则计算。模拟过程采用三维周期性边界条件,确保模拟晶胞至少扩展到 24 Å。其中,Lennard-Jones 势的截断半径设为 12 Å,并进行了长程校正,而 MOFs 和 H_2 分子之间的静电相互作用则通过 Ewald 求和法处理。

为达到体系平衡,我们在模拟中对吸附分子尝试平移、旋转、插入和删除等移动操作。每个 MOF 的 GCMC 模拟执行 1×10^6 周期,前 50000 个循环用于系统平衡,后 50000 个循环用于统计采样。每个循环中执行的移动尝试次数与该循环内体系中的吸附分子数相等。基于上述模拟结果,可进一步评估 MOFs 对 H_2 的吸附选择性。为了选取储氢性能最优异 MOFs,本次研究以工作容量(ΔW : 159K、30bar 和 298K、1bar 两个条件下 MOFs 的吸附量之差)为评价指标进行排序,将 MOFs 数据库划分为 TOP5%、TOP10%、TOP25%、TOP50% 和 All MOFs 五个级别。

1.3 机器学习

作为一种能够通过不同算法从大数据中学习的强大计算工具,ML 除了能预测未知材料的性能外,还有助于识别与应用输入描述符与材料输出性能之间的隐藏关系。本研究采用 5 种 ML 算法对 7 个特征描述符进行训练和预测,其中 5 种 ML 算法分别是分类提升算法(Categorical Boosting, CatBoost)、随机森林(Random Forest, RF)、极限梯度提升回归器(Extreme Gradient Boosting Regression, XGBR)、极端随机树(Extremely Randomized Trees, ET)、轻量级梯度提升机(Light Gradient Boosting Machine, LGBM),而 7 种特征描述符则分别是 VSA、 φ 、LCD、

PLD、 ρ 、 Q_a 和 K 。为了提高 ML 预测的合理性,采用了五折交叉验证法对每个参数的各种可能值进行评估,优化关键参数后将所有数据分成 10 份,对于每个周期,随机选择 70% 的数据作为训练集,30% 的数据作为测试集。为确保模型的稳定性,每个算法重复运行 5 次,选取预测效果最好的结果为最终结果。本次模型精度采用平均决定系数(R^2)、平均绝对误差(MAE)以及均方根误差(RMSE)三个指标进行评估。

经过三个指标的综合评估,CatBoost 算法的预测效果最好,故在进一步的数据挖掘中,我们采用 CatBoost 模型结合 SHAP 算法深入分析特征的相对重要性。在本次研究中,分别对不同层级的 MOFs 依次进行了 SHAP 技术分析,从平均绝对 SHAP 值得出的全局特征重要性,进而量化每个特征对该样本预测结果的贡献度。通过分析所有样本的 SHAP 值,最终可解读模型的整体行为特征。

1.4 分子指纹

分子指纹(MF)是一种将分子拆解为子结构并将这些子结构转换为数值表示的技术。为设计出高性能的储氢吸附材料,本研究将 ML 与 MF 相结合,进而筛选出优异的子结构,这些子结构可能是设计高性能 MOF 材料的关键。首先,利用 OpenBabel^[40]软件包将 MOF 的“cif”格式转换为“sdf”格式。然后基于 PaDEL-descriptor^[41]软件包,将“sdf”格式的 MOF 结构转化为 MACCS 材料指纹^[42]。

MACCS 是一种基于二维描述符的分子指纹,长度为 166 位,每一位都对应着明确的化学含义,具有良好的化学可解释性。TAN 等人^[43]在预测 MOF 对碘的吸附性能研究中,于同一数据集上系统比较了四种分子指纹(MACCS, PubChem, AtomPairs2D, Estate)的预测性能,结果表明 MACCS 指纹的预测精度最高,优于其他三种指纹。此外,现有研究表明^[44, 45],在利用机器学习预测 MOF 气体吸附性能的研究领域中,采用 MACCS 指纹结合 SHAP 技术分析可以有效识别出影响 MOF 吸附性能的关键子结构。

2 实验结果与讨论

2.1 低温储氢 MOF 的构效关系

本研究采用单变量分析探讨 MOFs 中不同描述符与 ΔW 之间的关系,进一步揭示了 MOFs 中的结构—性能关系,具体信息参见图 S6。如图 1a 所示, ΔW

随着 φ 的增大而增大,这主要是因为较高的 φ 能为 H_2 分子提供更多的附着位点和储存空间。然而当 $\varphi > 0.7$ 时, ΔW 的增速会趋于平缓,这说明 φ 超过某个临界值后,继续调整 φ 对优化 MOFs 吸附性能的效果不大。如图 1b 所示, $Q_{\text{st}} > 15\text{kJ}\cdot\text{mol}^{-1}$ 时,MOFs 的吸附性能较好,而当 $Q_{\text{st}} < 5\text{kJ}\cdot\text{mol}^{-1}$,MOFs 几乎不发生吸附行为,这是因为过低的吸附热使得 MOFs 和 H_2 分子之间的相互作用力很弱,不利于分子吸附。如图 1c、1d 所示, ΔW 随着 PLD、LCD 二者的变化趋势几乎是一致的,当 PLD、LCD 处于 10~20Å 时, ΔW 的数值较高,这主要是因为适中的 PLD、LCD 可以保证较高吸附速率的同时不影响 H_2 分子的解析过程。如图 1d 所示, ΔW 随着 ρ 的增大而减小,当 $\rho < 1000\text{kg}\cdot\text{m}^{-3}$ 时,MOFs 的 ΔW 处于较高水平,这是因为低 ρ 值常常意味着较高的 φ 值,而由前面的分析可得,高 φ 才是提升 MOFs 性能的关键参数。

2.2 复杂多变量分析——机器学习

单变量构效分析得到的是单一描述符对 MOFs 吸附性能的影响,而 ML 技术可以综合考虑多个描述符对 MOFs 吸附性能的影响,即通过多变量分析揭示 MOFs 结构-性能之间的关系。本次研究运用五种 ML 算法(RF, CatBoost, XGBR, LGBM, ET)训练和预测 MOFs 对氢气的吸附性能,在进行机器学习任务时,按照 7:3 的比例将原始数据分为训练集和测试集,其中训练集主要是用于训练和优化模型,而测试集则用于评估模型性能的好坏。算法的最优参数通过五折交叉验证与随机搜索确定,具体细节参见表 S2。如图 2a、2b、2c 及图 S8 所示,综合五种 ML 算法对其预测效果进行排序: CatBoost > XGBR > LGBM > RF > ET。其中 CatBoost 算法预测得到的测试集 R^2 高达 0.951,在预测 H_2 吸附性能上表现出卓越性。如图 2d 所示, CatBoost 算法预测得到的 MAE、RMSE 分别为 0.042 和 0.071,进一步验证了 CatBoost 算法的高精度和稳定性。与 XGBR、LGBM 算法不同的是, CatBoost 算法是基于对称决策树构建的。在每一步中,前一棵树的叶子都使用相同的条件进行拆分,选择损失最低的特征分割对并将其用于所有级别的节点。这种结构简化了模型,有助于提高预测速度,并作为一种正则化手段减少过拟合风险,进而提高了算法的准确度和泛化能力。

为了进一步研究不同描述符的相对重要性(RI),本次研究采用预测效果最好的 CatBoost 算法

对 5 个数据集进行了系统分析。如图 2e 所示,随着评估数据集向最优 MOFs 移动, φ 的 RI 逐渐下降, Q_{st} 的 RI 则是先上升后略微下降,而 ρ 的 RI 则在总体上呈现上升趋势。在 TOP5% 的数据集中,RI 排在前三位的分别是 Q_{st} 、 ρ 、 φ ,这说明从中等性能向高性能过渡时, Q_{st} 与吸附性能之间的关联性逐渐增强,即 Q_{st} 成为区分材料性能的关键因素。而在 TOP5% 的最优 MOFs 中, Q_{st} 的 RI 略有下降,则表明当材料的性能达到极致时, Q_{st} 可能已经达到了最优区间,不再是导致材料性能差异的主要原因。结合图 2e 中 MAACS 指纹的 RI 逐渐上升,说明在最优 MOFs 中 MACCS 指纹对 ΔW 的贡献度愈发显著。

2.3 MACCS 材料指纹与 SHAP 技术分析

MACCS 指纹共包含 166 个结构位点,为深入探究对吸附性能具有决定性影响的关键子结构,本研究采用 SHAP 技术对基于最优的 CatBoost 算法构建的模型进行解析,量化了各个指纹序列对吸附性能的贡献。本次研究对 TOP5%、TOP10%、TOP25%、TOP50%、ALL MOFs 五个数据集都进行了 SHAP 技术分析,筛选出在 MACCS 指纹中 RI 排名前 14 的分子指纹,分析不同分子指纹对 ΔW 的影响效果,具体细节参见图 S10。由图 S10 可知,第 99 号指纹在 5 个数据集中的 RI 都排在前列,这说明该指纹是构成 MOFs 框架的核心基础子结构。由表 1 可知,第 99 号指纹代表的子结构是碳碳双键,与自由旋转的碳碳单键不同,双键不能进行转动,这种结构能够为 MOFs 提供一个刚性框架。然而这只是一个 MOFs 的普适性指纹,想要开发出高性能的储氢吸附材料,需要进一步探究出助力性能优化的分子指纹。

如图 3 所示,在 TOP5% 的数据集中,第 99、62、45、65、97、66、158、137、148 号指纹对 ΔW 起促进作用,而 89、76、117、50、150 号指纹则对 ΔW 起抑制作用,其中第 99、62、89、45、65 号指纹是对 ΔW 影响最大的 5 个指纹。值得注意的是,第 99、62、89、45 四个指纹都是 TOP10% 数据集中的前 14 个最优指纹,而第 65 号指纹则不是其中之一,甚至第 65 号指纹也不是其他数据集中的前 14 个最优指纹。基于此可以进行合理猜测,第 99、62、89、45 号指纹可以适当提高 MOFs 的吸附性能,然而想让 MOFs 的吸附性能达到极致,则需要重点关注第 65 号指纹。图 S12 为第 65 号指纹对 CatBoost 模型预测 $\lg\Delta W$ 的 SHAP 依赖性图,由该图可知,当 MACCSFP65 的取值为 1,即 MOF 中存在该子结构时,样本对应的 SHAP 值大

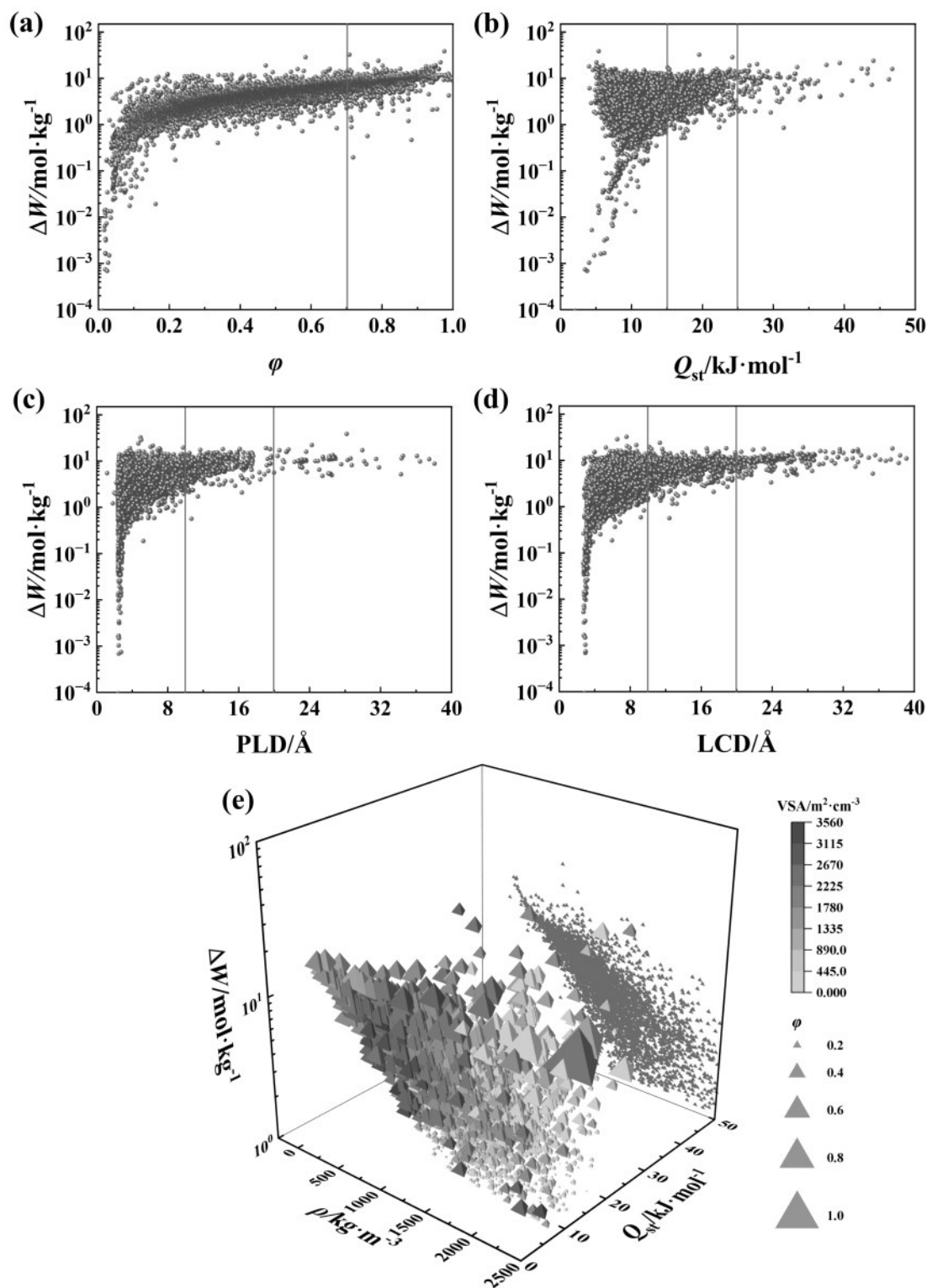


图1 MOFs特征描述符与吸附性能的关系: (a) ϕ - ΔW ; (b) Q_{st} - ΔW ; (c) PLD- ΔW ; (d) LCD- ΔW ; (e) ρ , Q_{st} , VSA, ϕ - ΔW ;

Fig.1 Relationship between Characteristic Descriptors of MOFs and their Adsorption Properties: (a) ϕ - ΔW ; (b) Q_{st} - ΔW ; (c) PLD- ΔW ; (d) LCD- ΔW ; (e) ρ , Q_{st} , VSA, ϕ - ΔW

多为正值且数值较高,此时MOF表现出较高的工作容量;而当该位点取值为0,即MOF中不存在该子结构时,SHAP值则集中于0附近或呈现负值,相应

MOF的工作容量也处于较低水平。上述结果表明,第65号指纹所对应的子结构对MOF的低温储氢性能具有显著的正向促进作用。由表1可知,第65号

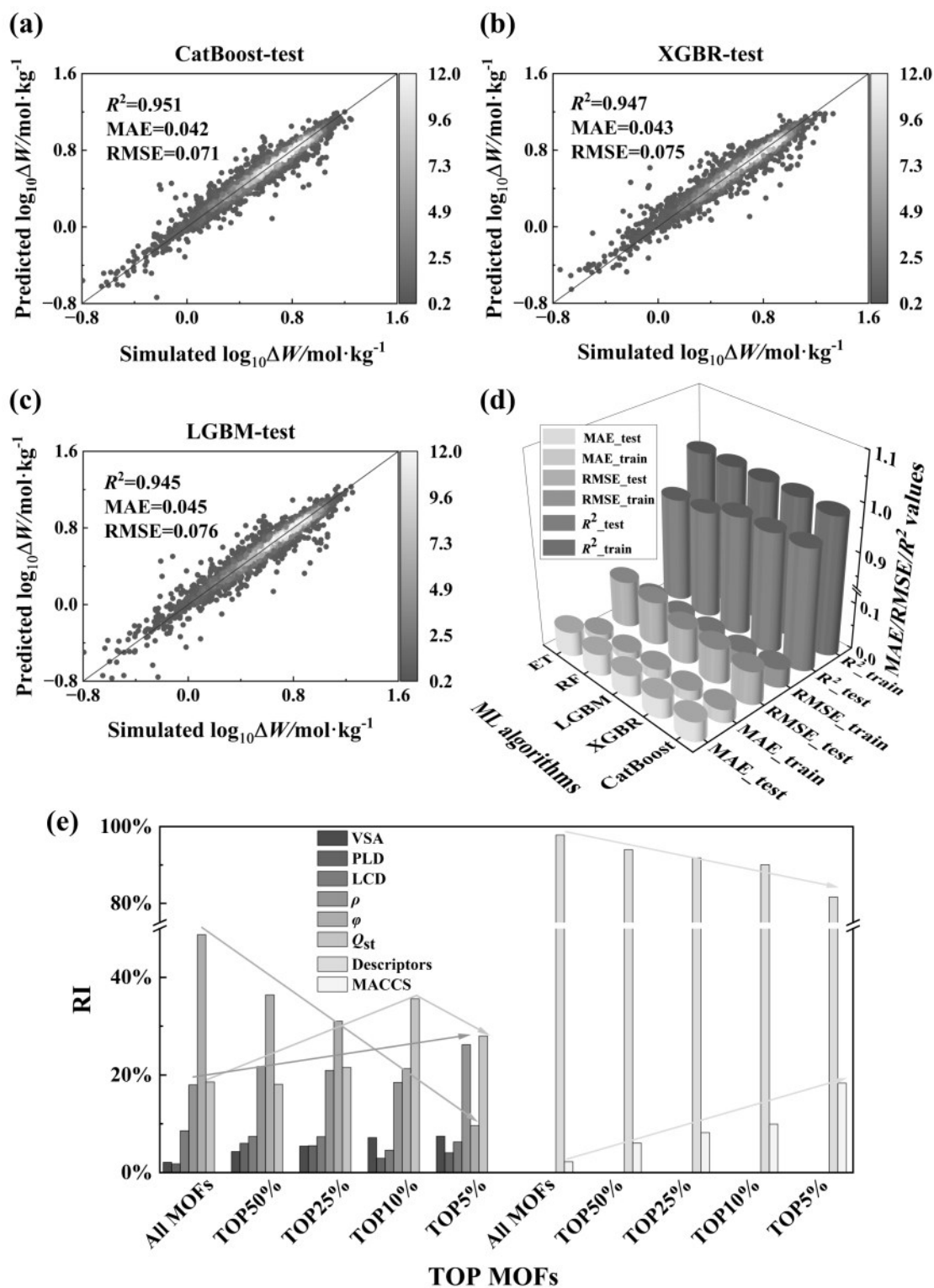


图2 (a) CatBoost, (b) XGBR 和(c) LGBM 三种 ML 算法预测 ΔW 性能对比; (d) 5 种 ML 算法在训练集和测试集上预测 ΔW 的性能指标; (e) 基于 CatBoost 算法预测不同特征描述符对 MOFs 的相对重要性

Fig.2 Comparison of ΔW Prediction Performance by (a) CatBoost, (b) XGBR and (c) LGBM Algorithms; (d) Performance Metrics of Five ML Algorithms on Training and Test Sets for ΔW Prediction; (e) Relative Importance of MOFs Feature Descriptors Based on the CatBoost Model

4a、4b所示。对于QOZQIJ, H_2 分子的附着位点主要集中在金属中心附近,而对于有着类似孔道结构的RUTNOK01, H_2 分子则主要吸附在金属中心以及咪唑有机链接上。该研究结果表明,咪唑基团可以增加MOFs中 H_2 分子的附着位点,进而提升MOFs在低温下对 H_2 的吸附性能。该材料指纹对提升MOFs吸附性能的独特作用,与其结构密不可分。咪唑基团的刚性结构能够增强MOFs框架的稳定性,形成较为稳定的孔道结构^[49,50],为 H_2 分子的吸附提供合适的吸附空间。

此外,现有的研究结果表明^[51-53], H_2 诱导偶极与芳香碳氮杂环电子云之间存在相互作用。为了进一步探索该静电相互作用对MOFs低温储氢性能的作用效果,本次研究绘制了QOZQIJ与RUTNOK01框架中H原子相对于Zn等原子的径向分布函数(Radial Distribution Function, RDF),如图4c、4d所示。由图4c可知,QOZQIJ框架中代表Zn-H相互作用的黑色曲线以及O-H相互作用的灰色曲线分别在3.99Å、3.75Å处出现了明显的特征峰,这表明 H_2 分子与MOF骨架中的开放金属位点(Zn)及周围的C-O结构存在一定的静电相互作用。而如图4d所示,RUTNOK01框架中代表N-H相互作用的灰色曲线在3.99Å处出现了一个显著的峰值,这表明H原子与咪唑基团中的N原子之间存在一种类似氢键的相互作用力,咪唑环与 H_2 的类似作用力在先前的工作中也被证明^[54],可以增加MOFs在低温下对 H_2 的吸附容量。且值得注意的是,RUTNOK01框架中代表Zn-H相互作用的黑色曲线在3.49Å处出现了较为尖锐的第一峰,该作用力大于QOZQIJ中的Zn-H相互作用力,这表明N-H之间存在的“类氢键”作用也在一定程度上增强了Zn-H之间的吸引力,类氢键和金属的协同增益作用也与BAG等人^[54]的工作结论相似,有助于MOFs低温储氢吸附性能的进一步提升。

3 结论

本研究通过系统筛选数万种CoRE-MOF材料,结合单变量分析揭示了不同结构/能量描述符与MOFs低温储氢吸附性能之间的构效关系。分析结果表明, Q_{st} 、 ρ 、 φ 三个描述符对于提高材料吸附性能至关重要,尤其是当 $Q_{st} > 15\text{kJ}\cdot\text{mol}^{-1}$ 、 $\rho < 1000\text{kg}\cdot\text{m}^{-3}$ 和 $\varphi > 0.7$ 时,MOFs的吸附性能处于较高

水平。随后采用5种ML算法训练和预测MOFs对 H_2 的低温吸附性能,其中CatBoost算法展现出最优异的预测效果。根据不同数据集中RI的变化趋势分析,当材料的性能达到TOP20%或TOP10%时, Q_{st} 等描述符对MOFs吸附性能的影响逐渐下降,而材料指纹对 ΔW 的贡献度则愈发显著。因此,本研究采用SHAP技术对基于CatBoost算法构建的模型进行解析,量化各个材料指纹序列对低温储氢性能(ΔW)的贡献。研究结果表明,在材料达到TOP10%之前,第99、62、45等指纹都可以有效提高MOFs对 H_2 的吸附作用;但在TOP5%的材料指纹中,大数据挖掘发现第65号指纹对性能的贡献度显著提高,从吸附密度图和RDF分析可知,在低温时,第65号指纹对应的芳香碳氮杂环结构可与氢气形成轻微的“类氢键”效果,使得氢气吸附量有所突破(从TOP10%上升到TOP5%)。因此,第65号指纹的引入是实现MOFs低温储氢吸附性能再上一个台阶的关键所在。基于大数据分析机器学习模型的系统融合,本次研究精准识别了提升MOFs低温储氢性能的核心子结构,证实其突破性能瓶颈的关键在于类氢键作用力的协同贡献,为高性能低温储氢吸附材料的定向设计提供了新的理论指导和设计思路。

符号说明

ΔW ——工作容量, $\text{mol}\cdot\text{kg}^{-1}$

$g(r)$ ——径向分布函数

K ——亨利系数, $\text{mol}\cdot\text{kg}^{-1}\cdot\text{Pa}^{-1}$

LCD——最大空腔直径, Å

PLD——孔径限制直径, Å

Q_{st} ——吸附热, $\text{kJ}\cdot\text{mol}^{-1}$

RI——相对重要性

VSA——体积比表面积, $\text{m}^2\cdot\text{cm}^{-3}$

q_i, q_j ——分别是原子*i*和*j*的原子电荷, *e*

r_{ij} ——原子对的距离, Å

u ——原子对之间的非键相互作用势能, $\text{kJ}\cdot\text{mol}^{-1}$

δ_{ij} ——原子相互碰撞时的最小距离, Å

ϵ_0 ——真空介电常数, $\text{C}^2\cdot\text{N}^{-1}\cdot\text{m}^{-2}$

ϵ_{ij} ——能量最低点到能量为零的差, $\text{kcal}\cdot\text{mol}^{-1}$

ρ ——密度, $\text{kg}\cdot\text{m}^{-3}$

φ ——孔隙率

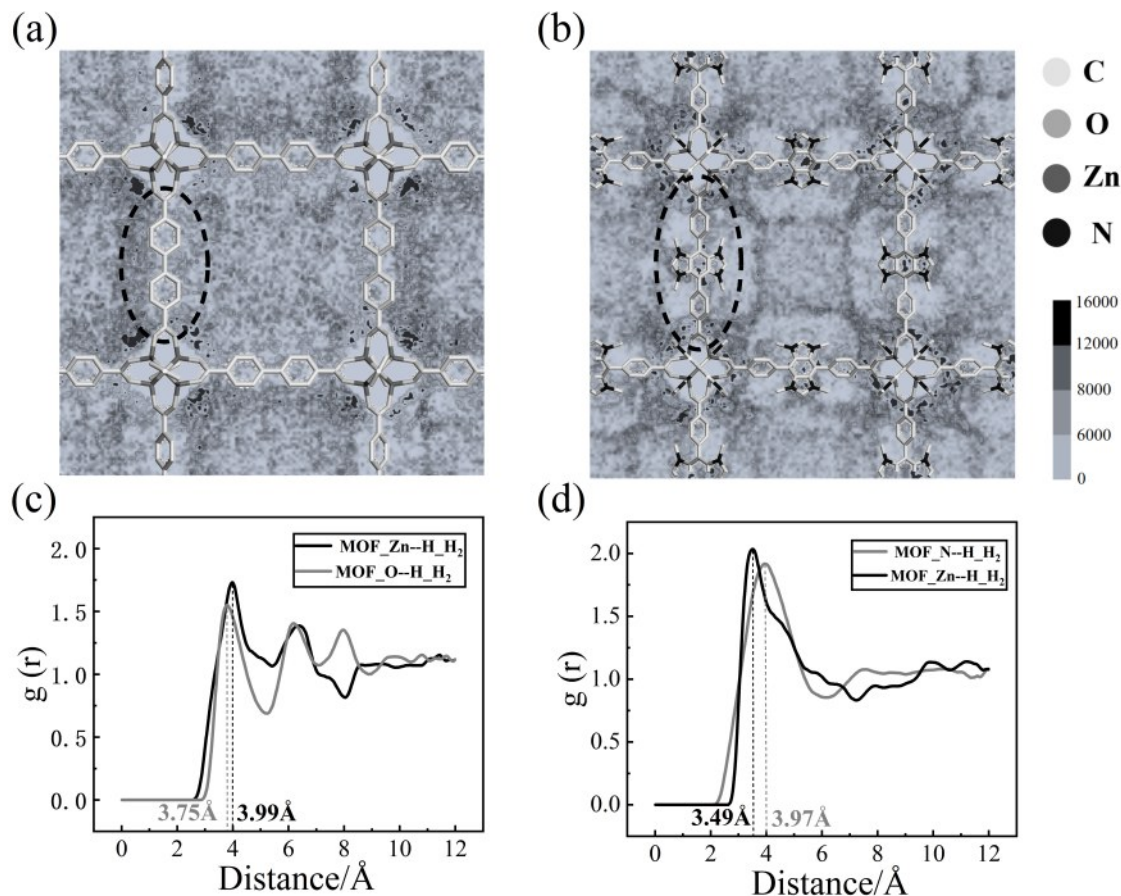


图4 (a) QOZQIJ, (b) RUTNOK1 在 159K、30bar 条件下对 H_2 的吸附密度图; (c) QOZQIJ 框架中 H_2 的 H 原子相对于 Zn 原子和 O 原子的径向分布函数; (d) RUTNOK1 框架中 H_2 的 H 原子相对于 Zn 原子和 N 原子的径向分布函数

Fig.4 Adsorption density maps of hydrogen by (a) QOZQIJ and (b) RUTNOK1 at 159K and 30 bar; (c) RDF of H atoms in H_2 with respect to Zn and O atoms in the QOZQIJ framework; (d) RDF of H atoms in H_2 respect to Zn and N atoms in the RUTNOK1 framework

参考文献

- [1] 赵赏鑫. "双碳" 目标下氢能产业高质量发展的管输战略与关键技术突破[J]. 油气储运, 2026, **45**(1): 1-11.
Zhao S X. Pipeline transportation strategy and key technological breakthroughs for the high-quality development of the hydrogen energy industry under the "dual carbon" goals[J]. Oil & Gas Storage and Transportation, 2026, **45**(1): 1-11.
- [2] Odenweller A, Ueckerdt F. The green hydrogen ambition and implementation gap[J]. Nature Energy, 2025, **10**(1): 110-123.
- [3] 刘佳育, 罗纳, 李俊辰, 等. 氢能航空发动机研究现状及发展制约因素[J]. 材料研究与应用, 2024, **18**(2): 299-308.
Liu J Y, Luo N, Li J C, et al. Research status and development constraints of hydrogen aero engines[J]. Materials Research and Application, 2024, **18**(2): 299-308.
- [4] 屈少鹏, 张海强, 杨璐嘉, 等. 海上风电制氢输运材料的研发现状及趋势[J]. 材料导报, 2026, **40**(2): 5-15.
Qu S P, Zhang H Q, Yang L J, et al. Research status and development trends of transport materials for offshore wind to hydrogen[J]. Materials Reports, 2026, **40**(2): 5-15.
- [5] 杨卫平, 杨浩坤, 曹树岭, 等. 面向氢能安全领域的抗氢脆高
强钢研究进展[J]. 材料研究与应用, 2025, **19**(5): 983-991, 10005.
Yang W P, Yang H K, Cao S L, et al. Research progress on hydrogen embrittlement-resistant high-strength steel for hydrogen energy safety[J]. Materials Research and Application, 2025, **19**(5): 983-991, 10005.
- [6] Zhou D S, Zheng C L, Zhang Y H, et al. An overview of RE-Mg-based alloys for hydrogen storage: Structure, properties, progresses and perspectives[J]. Journal of Magnesium and Alloys, 2025, **13**(1): 41-70.
- [7] Kamran M, Turzyński M. Exploring hydrogen energy systems: a comprehensive review of technologies, applications, prevailing trends, and associated challenges[J]. Journal of Energy Storage, 2024, **96**: 112601.
- [8] Hui Y Z, Wang M T, Guo S R, et al. Comprehensive review of development and applications of hydrogen energy technologies in China for carbon neutrality: Technology advances and challenges [J]. Energy Conversion and Management, 2024, **315**: 118776.
- [9] Wang X Y, Peng P, Witman M D, et al. Technoeconomic insights into metal hydrides for stationary hydrogen storage[J]. Advanced Science, 2025, **12**(21): 2415736.
- [10] Yang Z X, Wang Y Z, Lin X, et al. Vanadium induces Ni-Co

- MOF formation from a NiCo LDH to catalytically enhance the MgH_2 hydrogen storage performance[J]. *Journal of Magnesium and Alloys*, 2025, **13**(8): 4020–4031.
- [11] Ge L X, Qiu M H, Zhu Y F, et al. Synergistic catalysis of Ru single-atoms and zeolite boosts high-efficiency hydrogen storage[J]. *Applied Catalysis B: Environmental*, 2022, **319**: 121958.
- [12] Morandé A, Lillo P, Blanco E, et al. Modification of a commercial activated carbon with nitrogen and boron: Hydrogen storage application[J]. *Journal of Energy Storage*, 2023, **64**: 107193.
- [13] Moghadam P Z, Chung Y G, Snurr R Q. Progress toward the computational discovery of new metal-organic framework adsorbents for energy applications[J]. *Nature Energy*, 2024, **9**(2): 121–133.
- [14] Alezi D, Belmabkhout Y, Suyetin M, et al. MOF crystal chemistry paving the way to gas storage needs: aluminum-based soc-MOF for CH_4 , O_2 , and CO_2 storage[J]. *Journal of the American Chemical Society*, 2015, **137**(41): 13308–13318.
- [15] Connolly B M, Madden D G, Wheatley A E H, et al. Shaping the future of fuel: monolithic metal-organic frameworks for high-density gas storage[J]. *Journal of the American Chemical Society*, 2020, **142**(19): 8541–8549.
- [16] Wang P F, Teng Y, Zhu J L, et al. Review on the synergistic effect between metal-organic frameworks and gas hydrates for CH_4 storage and CO_2 separation applications[J]. *Renewable and Sustainable Energy Reviews*, 2022, **167**: 112807.
- [17] Hu A Q, Xie Q Y, Chen L Y, et al. Hierarchically ordered meso-/macroporous MOF-based materials for catalysis and energy applications[J]. *EnergyChem*, 2024, **6**(6): 100137.
- [18] Wang Y, Jiang Z J, Lu W G, et al. Machine learning-assisted exploration of chemical space of MOF-5 analogs for enhanced $\text{C}_2\text{H}_6/\text{C}_2\text{H}_4$ separation[J]. *Angewandte Chemie International Edition*, 2025, **64**(21): e202500783.
- [19] Ma X, Sun N, Li Z G, et al. Highly flexible and self-standing covalent organic framework-metal-organic framework (COF-MOF) composite crystalline porous material (CPM) membrane for molecular separation[J]. *Advanced Functional Materials*, 2024, **34**(19): 2312203.
- [20] Cong S Z, Zhou Y Z, Luo C L, et al. Designing metal-organic framework (MOF) membranes for isomer separation[J]. *Angewandte Chemie International Edition*, 2024, **63**(15): e202319894.
- [21] Liu S Y, Zhang Y, Zhu F Z, et al. Mg-MOF-74 derived defective framework for hydrogen storage at above-ambient temperature assisted by Pt catalyst[J]. *Advanced Science*, 2024, **11**(18): 2401868.
- [22] Chen J, Tang Z Z, Zhu D, et al. Charge-induced defective MOF-801 with enhanced sorption heat for efficient hydrogen storage[J]. *Small*, 2025, **21**(23): 2502860.
- [23] Zhao H Q, Zhang H M, Wang J S, et al. Theoretical calculation and tests of hydrogen storage properties of Nano MoS_2 doped Ce-MOF-808[J]. *Journal of Energy Storage*, 2024, **88**: 111526.
- [24] Moghadam P Z, Li A, Wiggan S B, et al. Development of a Cambridge structural database subset: a collection of metal-organic frameworks for past, present, and future[J]. *Chemistry of Materials*, 2017, **29**(7): 2618–2625.
- [25] Han Z S, Yang Y H, Rushlow J, et al. Development of the design and synthesis of metal-organic frameworks (MOFs)-from large scale attempts, functional oriented modifications, to artificial intelligence (AI) predictions[J]. *Chemical Society Reviews*, 2025, **54**(1): 367–395.
- [26] Daglar H, Gulbalkan H C, Avci G, et al. Effect of metal-organic framework (MOF) database selection on the assessment of gas storage and separation potentials of MOFs[J]. *Angewandte Chemie International Edition*, 2021, **60**(14): 7828–37.
- [27] 程洪, 葛美伶, 司天宇, 等. 机器学习辅助金属材料力学性能预测[J]. *材料研究与应用*, 2023, **17**(6): 1070–1077.
- Cheng H, Ge M L, Si T Y, et al. Machine learning-assisted prediction of mechanical properties of metallic materials[J]. *Materials Research and Application*, 2023, **17**(6): 1070–1077.
- [28] Jang J, Gu G H, Noh J, et al. Structure-based synthesizability prediction of crystals using partially supervised learning[J]. *Journal of the American Chemical Society*, 2020, **142**(44): 18836–18843.
- [29] Kim W T, Lee W G, An H E, et al. Machine learning-assisted design of metal-organic frameworks for hydrogen storage: a high-throughput screening and experimental approach[J]. *Chemical Engineering Journal*, 2025, **507**: 160766.
- [30] Wang X Y, Breunig H M, Peng P. Broad range material-to-system screening of metal-organic frameworks for hydrogen storage using machine learning[J]. *Applied Energy*, 2025, **383**: 125346.
- [31] 胡嘉朗, 姜明源, 金律铭, 等. 机器学习辅助MOFs高通量计算筛选及气体分离研究进展[J]. *化工学报*, 2025, **76**(5): 1973–1996.
- Hu J L, Jiang M Y, Jin L M, et al. Machine learning-assisted high-throughput computational screening of MOFs and advances in gas separation research[J]. *CIESC Journal*, 2025, **76**(5): 1973–1996.
- [32] 高正, 汪辉, 屈治国. 数据驱动辅助高通量筛选阴离子柱撑金属有机框架储氢[J]. *化工学报*, 2025, **76**(8): 4259–4272.
- Gao Z, Wang H, Qu Z G. Data-driven high-throughput screening of anion-pillared metal-organic frameworks for hydrogen storage[J]. *CIESC Journal*, 2025, **76**(8): 4259–4272.
- [33] Chung Y G, Haldoupis E, Bucior B J, et al. Advances, updates, and analytics for the computation-ready, experimental metal-organic framework database: core MOF 2019[J]. *Journal of Chemical & Engineering Data*, 2019, **64**(12): 5985–5998.
- [34] Dubbeldam D, Calero S, Ellis D E, et al. RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials[J]. *Molecular Simulation*, 2016, **42**(2): 81–101.
- [35] Rappe A K, Casewit C J, Colwell K S, et al. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations[J]. *Journal of the American Chemical Society*, 1992, **114**(25): 10024–10035.
- [36] Li Y, Situ Y, Guan K X, et al. High dynamic separation performance of metal-organic frameworks for D_2/H_2 : Independent or competitive adsorption[J]. *AIChE Journal*, 2024, **70**(1): e18283.
- [37] Darkrim F, Levesque D. Monte Carlo simulations of hydrogen adsorption in single-walled carbon nanotubes[J]. *The Journal of Chemical Physics*, 1998, **109**(12): 4981–4984.
- [38] Cao D W, Ren J H, Gong Y, et al. Quantum sieving of H_2/D_2 in MOFs: a study on the correlation between the separation performance, pore size and temperature[J]. *Journal of Materials*

- Chemistry A, 2020, **8**(13): 6319–6327.
- [39] Zhao F J, Tan Y X, Wang W J, et al. Optimizing H₂, D₂, and C₂H₂ sorption properties by tuning the pore apertures in metal–organic frameworks[J]. *Inorganic Chemistry*, 2018, **57**(21): 13312–7.
- [40] O’Boyle N M, Banck M, James C A, et al. Open Babel: an open chemical toolbox[J]. *Journal of Cheminformatics*, 2011, **3**(1): 33.
- [41] Yap C W. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints[J]. *Journal of Computational Chemistry*, 2011, **32**(7): 1466–1474.
- [42] Durant J L, Leland B A, Henry D R, et al. Reoptimization of MDL keys for use in drug discovery[J]. *Journal of Chemical Information and Computer Sciences*, 2002, **42**(6): 1273–1280.
- [43] Tan H Y, Teng Y K, Shan G C. High throughput computational screening and interpretable machine learning for iodine capture of metal–organic frameworks[J]. *npj Computational Materials*, 2025, **11**: 115.
- [44] Lin Z H, Cai H, Peng H J, et al. Enhancing arsenate removal through interpretable machine learning guiding the modular design of metal–organic frameworks[J]. *Chemical Engineering Journal*, 2024, **497**: 155058.
- [45] Ming Z Y, Zhang M, Zhang S X, et al. A multi-method approach to analyzing MOFs for chemical warfare simulant capture: molecular simulation, machine learning, and molecular fingerprints[J]. *Nanomaterials*, 2025, **15**(3): 183–183.
- [46] Xian J Y, Xie X X, Huang Z Y, et al. Structure and properties of a mixed-ligand co-MOF that was synthesized in situ from a single imidazole-pyridyl-tetrazole trifunctional ligand[J]. *Crystal Growth & Design*, 2023, **23**(3): 1448–1454.
- [47] Chen D M, Xu N, Qiu X H, et al. Functionalization of metal–organic framework via mixed-ligand strategy for selective CO₂ sorption at ambient conditions[J]. *Crystal Growth & Design*, 2015, **15**(2): 961–965.
- [48] Qiao Z W, Xu Q S, Jiang J W. Computational screening of hydrophobic metal–organic frameworks for the separation of H₂S and CO₂ from natural gas[J]. *Journal of Materials Chemistry A*, 2018, **6**(39): 18898–18905.
- [49] 翁惠琼, 黄河, 王雯菲, 等. 基于MOFid赋能下的AI大数据挖掘高性能化学战剂吸附材料[J]. *化学学报*, 2026, **84**(1): 8–19.
- Weng H Q, Huang H, Wang W F, et al. AI big-data mining empowered by MOFid for high-performance chemical warfare agent adsorbents[J]. *Acta Chimica Sinica*, 2026, **84**(1): 8–19.
- [50] Sengupta D, Melix P, Bose S, et al. Air-stable Cu(I) metal–organic framework for hydrogen storage[J]. *Journal of the American Chemical Society*, 2023, **145**(37): 20492–20502.
- [51] Nguyen-Thuy T, Le-Hoang P, Hoang Vu N, et al. Hydrogen adsorption mechanism of MOF-74 metal–organic frameworks: an insight from first principles calculations[J]. *RSC Advances*, 2020, **10**(72): 43940–43949.
- [52] Fan C B, Wei Y Y, Wang L L, et al. Design and construction of cobalt–organic framework with outstanding adsorption properties assembled by 3-(3, 5-dicarboxylphenoxy)pyridine[J]. *Journal of Molecular Structure*, 2023, **1288**: 135733.
- [53] Meng Z S, Li Q Y, Sun H L, et al. Tricycloquinazoline-based monolayer conjugated metal–organic frameworks as promising hydrogen storage media: a theoretical investigation[J]. *Green Energy & Environment*, 2025, **10**(6): 1326–1336.
- [54] Bag A, Sinha S, Das H S, et al. Hydrogen storage efficiency of the Ag (I)/Au (I) decorated five-member aromatic heterocyclic (AH) compounds: a theoretical investigation[J]. *Engineered Science*, 2023: 1062–1062.