

DOI:10.11918/202402004

基于机器学习的纳滤膜预测筛选模型构建与评估

徐达梁, 徐杭滨, 靳心瑶, 刘超, 费兆轩, 姚杰, 张子峰, 李圭白, 梁恒

(城市水资源与水环境国家重点实验室(哈尔滨工业大学), 哈尔滨 150090)

摘要: 纳滤净水技术是应对水资源危机和水质安全保障的核心技术之一。然而, 纳滤膜性能长期受渗透性与选择性制约, 亟需开发高性能纳滤膜。纳滤膜制备过程涉及水相单体质量分数、水相添加剂质量分数、油相单体质量分数、聚合时间等因素, 传统的试误实验法需消耗大量的人力、物力与财力。依据纳滤膜制备参数, 构建基于机器学习的纳滤膜预测筛选模型。结果表明, XGBoost 机器学习模型可有效预测纳滤膜纯水通量与截留性能, 对纯水通量和截留性能的 R^2 评价指标分别为 0.84 和 0.90。采用 SHAP 值法对 XGBoost 机器学习模型中的输入参数进行量化分析, 发现水相单体质量分数与基膜类型对纯水通量有最高的绝对平均 SHAP 值, 分别为 2.77 与 2.59, 而面向纳滤膜截留性能的关键参数绝对平均 SHAP 值相对接近。单体子结构特征分析结果显示, 亲水性子结构特征与支链型子结构特征有助于提升纳滤膜纯水通量, 胺基则促进纳滤膜的截留性能。构建的纳滤膜预测筛选模型有助于关键参数的识别与优化, 为纳滤膜的开发提供理论与技术指导。

关键词: 纳滤膜; 机器学习; 性能预测; XGBoost 模型

中图分类号: TU991

文献标志码: A

文章编号: 0367-6234(2024)06-0008-08

Construction and evaluation of prediction model for nanofiltration membranes based on machine learning

XU Daliang, XU Hangbin, JIN Xinyao, LIU Chao, FEI Zhaoxuan, YAO Jie, ZHANG Zifeng, LI Guibai, LIANG Heng

(State Key Laboratory of Urban Water Resources and Water Environment (Harbin Institute of Technology), Harbin 150090, China)

Abstract: Nanofiltration is one of the core technologies to address water crises and ensure water quality safety and security. However, the performance of nanofiltration membranes has long been limited by the permeance/selectivity trade-off. Thus, developing high-performance nanofiltration membranes is an urgent need. Nanofiltration membrane fabrication involves factors such as aqueous phase monomer concentration, aqueous phase additive concentration, oil phase monomer concentration, polymerization time, etc. Traditional trial-and-error experimental method consumes substantial manpower, material and financial resources. In this study, based on the fabrication parameters of nanofiltration membranes, we constructed a machine learning-based predictive screening model for nanofiltration membranes. The results show that the XGBoost machine learning model effectively predict the permeance and rejection performance of nanofiltration membranes, with R^2 evaluation scores of 0.84 and 0.90, respectively, for the permeance and rejection performance. The quantitative analysis of the input parameters in the XGBoost machine learning model using the SHAP value method reveals that the aqueous-phase monomer concentration and the substrate type had high average SHAP values of 2.77 and 2.59 for permeance. The average SHAP values of the key parameters oriented towards the rejection performance were relatively close. The results of the monomer substructure characterization show that the hydrophilic and branched substructure features contribute to the permeance, while the amine group promote the rejection performance. The nanofiltration membrane prediction and screening model can identify and optimize the key parameters, providing theoretical and technical support for developing nanofiltration membranes.

Keywords: nanofiltration membranes; machine learning; performance prediction; XGBoost model

纳滤以其高效分离、出水稳定、集成化高与自动化运维等技术优势^[1], 在地表水净化、污水深度处理、工业废水回用领域受到广泛的关注与研究^[2-4]。

然而, 纳滤膜性能长期受渗透性与选择性的制约^[5], 研发先进膜材料是推广纳滤膜应用的必由之路。由于膜制备过程具有关键参数多、数值分布广、

收稿日期: 2024-02-02; 录用日期: 2024-03-11; 网络首发日期: 2024-04-18

网络首发地址: <https://link.cnki.net/urlid/23.1235.T.20240416.1423.002>

基金项目: 国家自然科学基金(52300083); 哈尔滨工业大学城市水资源与水环境国家重点实验室开放基金(QA202323)

作者简介: 徐达梁(1993—), 男, 讲师; 梁恒(1979—), 男, 教授, 博士生导师

通信作者: 梁恒, hitliangheng@163.com

组合性多样的特征,难以通过传统实验试错法逐个分析^[6],亟需构建膜制备参数-膜分离效能逆向设计的新范式^[7]。

数据科学与材料研究相结合的机器学习方法具有加速膜材料研发的潜力^[8]。机器学习通过分析多维数据集中的潜在趋势与规律,对新数据集进行预测与筛选^[9]。因此,利用机器学习可建立膜制备参数-膜分离效能间的关联。部分机器学习方法已应用于预测膜的渗透性、分离性能与污染过程,包括神经网络、随机森林与支持向量机等^[10-12]。Gao等^[13]利用分类树模型明确影响超滤膜性能的3个主要制备参数,优化超滤膜制备条件,实现超滤膜纯水通量与蛋白分离的同步提升。然而,影响纳滤膜性能的关键制备参数仍不清晰,机器学习算法应用于纳滤膜制备过程的可行性有待研究。

本研究通过膜制备过程参数归集,构建基于梯度提升集成学习的纳滤膜预测筛选模型。采用统计学方法评估机器学习模型对纳滤膜性能的预测能力;使用SHAP值法解析影响纳滤膜性能的关键参数,分析关键参数对纳滤膜性能的影响机制,明确适用于纳滤膜材料的子结构特征,为高性能纳滤膜材料的开发与应用提供理论与技术支撑。

1 模拟和方法

1.1 数据集构建

基于近十年纳滤膜文献构建机器学习数据集,其中,纳滤膜纯水通量数据690个,截留数据1550个。输入参数包括水相单体类型、水相添加剂类型、水相单体质量分数、水相添加剂质量分数、油相单体类型、油相单体质量分数、有机溶液类型、聚合时间、热交联温度、热交联时间、基膜类型、基膜孔径等纳滤膜制备过程的典型条件。对于文献中缺少的参数记录为缺失值,最终形成了纳滤膜纯水通量与截留性能两个数据集。

1.2 数据预处理

使用摩根指纹指代不同的水相单体分子。对于分子中的每个原子,基于拓扑距离形成相应的子结构特征集合,并将子结构特征转换为二进制数据。随着拓扑距离设置值增大,一个分子求解的特征指纹量增加,导致位相量上摩根指纹重叠。为了避免上述情况,将摩根指纹的拓扑距离设置为0,与机器学习模型的超参数一同优化。针对类别型数据,使用反向差分编码与Helmert编码方式转化为特征编码。对于纯水通量的数值型数据,基于100 kPa的跨膜压差进行均一化处理,通过幂变换使数据呈现高斯分布;对于截留性能的数值型数据,使用鲁棒性

缩放法处理,避免数据集内异常值的影响。

1.3 机器学习模型

选择两种梯度提升集成学习算法:XGBoost与LightGBM机器学习模型。将80%的数据集划分为训练集,剩余20%的数据集划分为测试集,定义机器学习模型的超参数范围,以五折交叉验证法构建目标函数,采用贝叶斯算法搜索目标函数的最优超参数集。对于数据集中的缺失值,使用现有数据计算最佳分裂点及划分至左子树或右子树的增益,选择增益较大的方向作为节点中特征缺失样本的分配方向。

1.4 模型评价

通过结合决定系数(R^2)、平均绝对误差(E_{MA})与均方根误差(E_{RMS})评估机器学习模型的预测效果,相关公式如下:

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

$$E_{MA} = \frac{\sum_{i=1}^n |(\hat{y}_i - y_i)|}{n}$$

$$E_{RMS} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{N}}$$

式中: y_i 为第*i*个样本的实际值, \hat{y}_i 为模型对第*i*个样本的预测值, \bar{y} 为数据集中所有样本的平均值。

1.5 特征评估

SHAP是一种基于博弈论手段解释机器学习模型预测结果的方法,可用于模型的全局解释和局部解释。采用SHAP值法评估每个输入参数对于模型的重要性。SHAP值按照以下公式计算:

$$\phi(f, x) = \sum_{S \subseteq N} \frac{(|S| - 1)! (n - |S| - 1)!}{n!} (f(S) - f(S \setminus x))$$

式中: $\phi(f, x)$ 为参数*x*的SHAP值, N 为所有输入参数的集合, S 表示包含参数*x*的子集, $|S|$ 为*S*中的参数数量, n 为输入参数数量。

2 结果与讨论

2.1 数据集统计分析

高性能纳滤膜制备与应用的关键参数包括基膜结构与特性、界面聚合反应过程、后处理条件等。其中,水相单体、水相添加剂、油相单体的种类、质量分数、浸泡时间对纳滤膜性能有显著影响,图1为数据集中水相单体质量分数、水相添加剂质量分数、油相

单体质量分数与聚合时间的统计分析图。结果表明,水相单体质量分数主要集中在 0.35 ~ 2.00 (图 1(a)),当采用疏水性基膜、弱活性水相单体或低水相浸泡时间等条件时,需增加水相单体质量分数使基膜表面吸附足够的水相单体用于后续的界面聚合反应。水相添加剂质量分数的第一和第三四分位数为 0.10% 与 1.00% (图 1(b)),数据分布范围相比水相单体质量分数的分布范围更窄,且投加量通常低于水相单体质量分数。水相单体和油相单体是界面聚合形成有机高分子聚合物的主要物质,水

相添加剂可局部参与界面聚合反应,调控有机高分子聚合物的孔径、孔隙率或自由体积分数,过量添加将抑制有机高分子聚合物的形成,因此,投加量通常低于水相单体质量分数。油相单体质量分数主要集中在 0.01 ~ 0.50 (图 1(c)),数值分布较大。聚合时间主要分布在 5 ~ 240 s (图 1(d)),中位数为 60 s,数据分布集中,特征异常值少。综上,纳滤膜制备过程呈现关键参数多、数值分布广、组合性多样的特征,由于膜制备参数与膜分离效能间的关系不明,阻碍了高性能纳滤膜的开发。

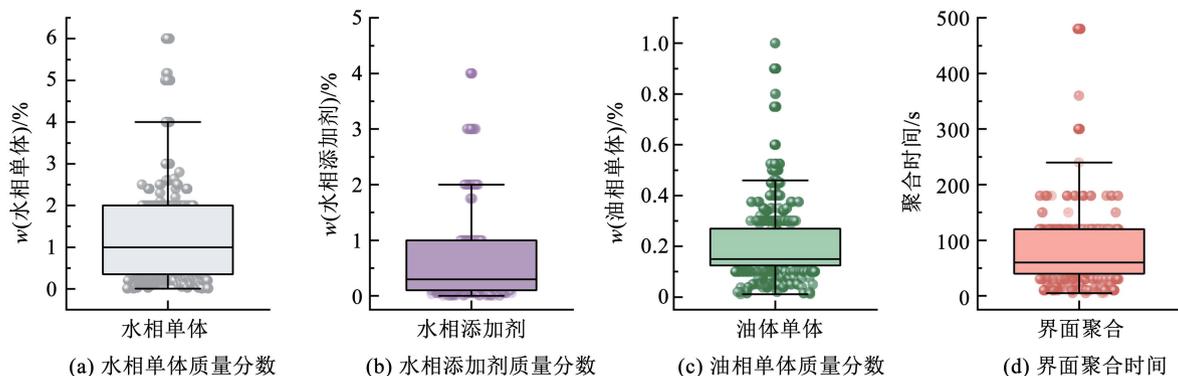


图 1 纳滤膜制备关键参数统计分析

Fig. 1 Statistical analysis of nanofiltration membrane fabrication parameters

2.2 模型预测能力评价

基于纳滤膜制备参数的数据集,采用机器学习模型预测相应纳滤膜纯水通量与截留性能。其中,80%的数据集样本随机划分为训练集,剩余 20% 的样本划分为测试集。在数据集处理过程中,利用摩根指纹指代不同的水相单体,通过反向差分编码与 Helmert 编码方式转化类别型数据,使用鲁棒性缩放法归一化数值型数据,降低数据集中异常值的影响。定义机器学习模型的超参数范围,以五折交叉验证法构建基于 E_{RMS} 的目标函数,通过贝叶斯算法搜索目标函数的最优超参数集,形成 XGBoost 与 LightGBM 机器学习模型,测试集与训练集的预测值和真实值对比如图 2 所示。总体而言,XGBoost 与 LightGBM 机器学习模型对纳滤膜纯水通量与截留性能都具有较好的预测能力,测试集中的数据大部分位于真实值的 95% 置信区间。XGBoost 与 LightGBM 均属于梯度提升集成学习算法,表明梯度提升决策树算法在纳滤膜性能预测方面具有优异的性能。Lu 等^[14]使用基于纳滤膜结构参数的数据集,通过 XGBoost 机器学习模型预测纳滤膜对一价/二价离子的选择性分离性能,结果显示,XGBoost 机器学习模型可有效预测纳滤膜的选择分离特性,并为纳滤膜结构参数的优化提供理论依据。然而,对于使用不同制备方法的二维层状膜、共价有机框架

膜、氢键有机框架膜,由于制备原理的不同及数据的有限性,难以通过当前的模型预测。

表 1 使用统计学指标进一步分析 XGBoost 与 LightGBM 机器学习模型对纯水通量和截留性能的预测能力。纯水通量预测方面,XGBoost 机器学习模型在训练集预测方面均优于 LightGBM 机器学习模型;二者在测试集上的 R^2 值一致,均为 0.84,略低于训练集 R^2 值;XGBoost 机器学习模型的测试集 E_{MA} (2.95) 略低于 LightGBM 机器学习模型 (3.05); XGBoost 机器学习模型的测试集 E_{RMS} 与 LightGBM 机器学习模型相近。上述结果表明,XGBoost 机器学习模型在纯水通量预测方面略优于 LightGBM 机器学习模型,但二者预测能力相近,均可依据纳滤膜制备参数有效预测纳滤膜的纯水通量。

截留率预测方面,XGBoost 机器学习模型在测试集上的性能均优于 LightGBM 机器学习模型, R^2 为 0.90、 E_{MA} 为 6.26、 E_{RMS} 为 9.18,表明 XGBoost 机器学习模型在处理高维稀疏数据时具有较好的效果,这与 XGBoost 算法采用二阶求导最小化目标损失有关。相比纯水通量预测,XGBoost 机器学习模型对截留率预测具有更高的 R^2 值,该结果是由于盐截留率数据集具有更大的数据量,提高了 XGBoost 机器学习模型的预测能力。综上,XGBoost 与 LightGBM 机器学习模型在测试集上都具有良好拟合精度,证

明了机器学习模型对纳滤膜性能预测的准确性。由于 XGBoost 机器学习模型在截留性能预测方面具有

更好的性能, 在后续研究中主要使用 XGBoost 机器学习模型进行分析。

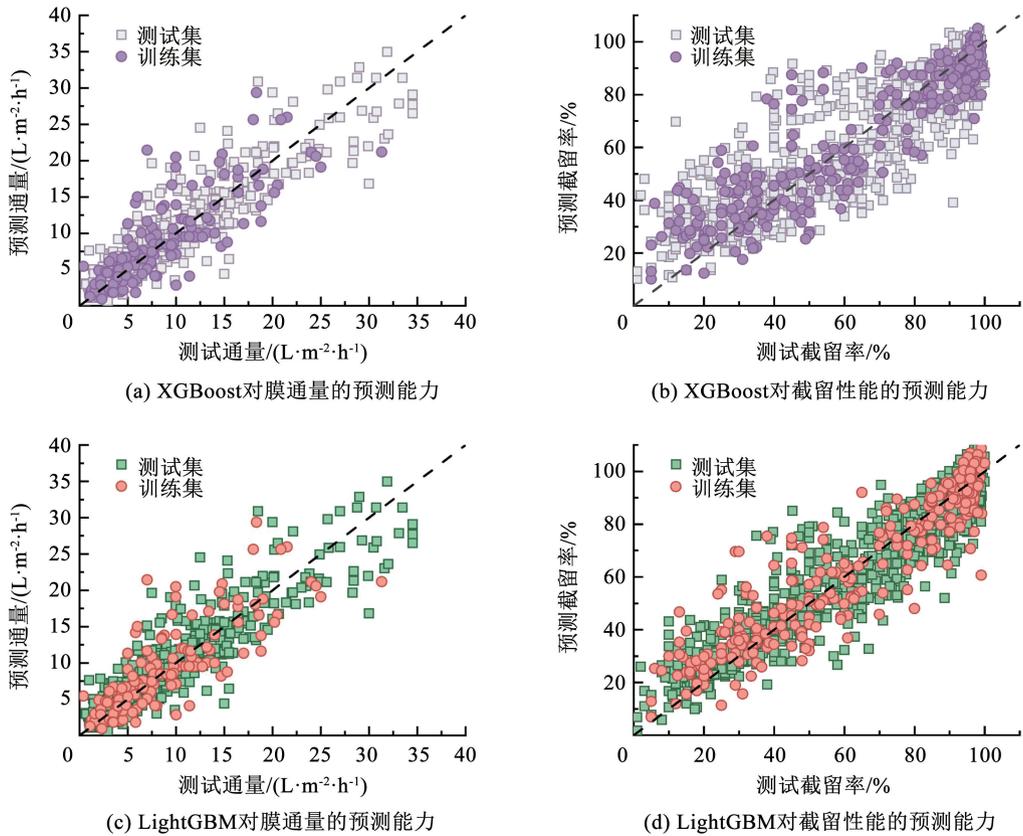


图2 机器学习对膜通量和截留性能的预测能力

Fig. 2 Prediction capability of machine learning models for membrane permeance and salt rejection

表1 机器学习对纯水通量和截留率的预测能力评价

Tab. 1 Prediction capability evaluation of machine learning models for pure water permeance and salt rejection

算法	目标对象	训练集 R^2	训练集 E_{MA}	训练集 E_{RMS}	测试集 R^2	测试集 E_{MA}	测试集 E_{RMS}
XGBoost	纯水通量	0.92	1.77	2.86	0.84	2.95	4.28
	截留率	0.98	2.76	3.10	0.90	6.26	9.18
LightGBM	纯水通量	0.89	2.22	3.39	0.84	3.05	4.30
	截留率	0.91	6.16	8.57	0.86	7.91	10.99

2.3 机器学习模型的解释分析

使用 SHAP 值法对机器学习模型进行解析, SHAP 值是对参数权重的无偏估计, 绝对平均 SHAP 值越大, 说明参数对于模型的影响越大^[15]。通过计算数据集中每个样本参数的绝对平均 SHAP 值, 可以明确关键制膜参数, 图 3 为基于 XGBoost 机器学习模型的膜制备参数绝对平均 SHAP 值。基于参数权重排序可以看出, 水相单体质量分数、基膜类型、运行压力、热交联时间、基膜膜孔、热交联温度、聚合时间对纳滤膜纯水通量与截留性能均有显著影响。

针对纳滤膜纯水通量, 水相单体质量分数与基膜类型具有最高的绝对平均 SHAP 值, 分别为 2.77 与 2.59。因此, 在高通量纳滤膜制备过程中, 需要

优先调控水相单体质量分数。同时, Karan 等^[16]通过调控基膜的化学结构特性, 改变水相单体的吸附容量与扩散速率, 优化界面聚合的反应过程, 成功制备出高通量纳滤膜。水相添加剂质量分数的绝对平均 SHAP 值为 1.31, 在前期研究中, 通过使用水相添加剂制备纳米复合纳滤膜, 可以在维持或小幅度降低截留率的情况下显著提升纯水通量^[17]。运行压力、油相单体质量分数、热交联时间等参数对纳滤膜纯水通量均有一定影响。

通过纳滤膜截留性能可以看出, 10 种关键参数的绝对平均 SHAP 值接近。聚合时间为纳滤膜截留性能最重要的影响因素, 绝对平均 SHAP 值为 2.07。基膜类型为次要影响因素, 表明纳滤膜制备过程

中不仅需要关注界面聚合的反应参数,也需选择合适的基膜作为反应载体,并提供足够的机械稳定性。水相单体的子结构特征 2633 和 2641 均起到一定的

影响,绝对平均 SHAP 值为 1.89 和 1.07。因此,在调控纳滤膜截留性能时需关注水相单体的子结构特征。

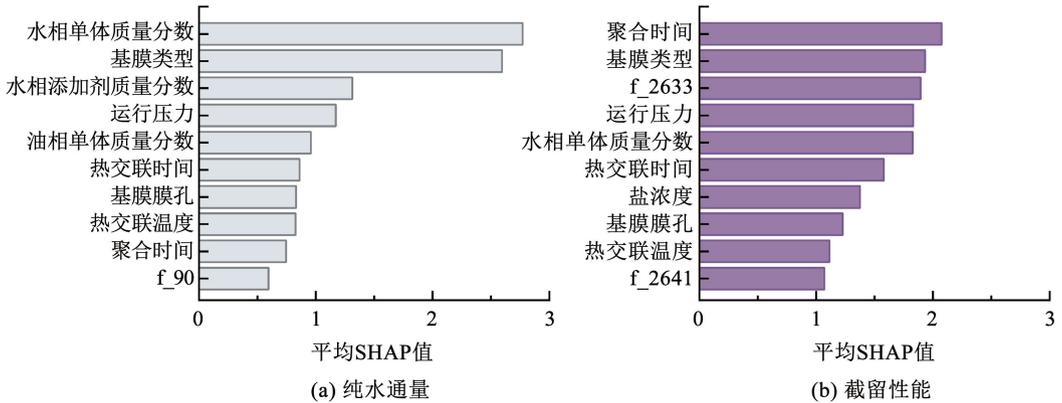


图 3 膜制备参数平均 SHAP 值

Fig. 3 Mean SHAP values of membrane fabrication parameters

基于膜制备参数的权重排序,分析数据集内关键参数对纳滤膜纯水通量与截留性能的影响,具体结果如图 4 所示。各参数的 SHAP 值分布于特征线上,原点右侧的点表示该参数对目标结果的正向作用,原点左侧的点则表示该参数对目标结果有反向作用,且 SHAP 值的绝对值越大对目标结果的影响越大^[18]。纯水通量(图 4(a))结果显示,高水相单体质量分数将降低纳滤膜通量,这是由于水相单体质量分数升高将增加纳滤膜的交联度,导致纳滤膜通量下降。由于基膜类型是类别型参数,难以通过 SHAP 值直观判断对纳滤膜性能的影响,水相单体的子结构特征 90 对纳滤膜通量有反向作用,可以适当降低水相单体中的子结构特征 90,实现高通量纳

滤膜的开发。

纳滤膜截留性能(图 4(b))表明,较高的聚合时间对截留性能有负面影响,上述结果与实际反应过程相悖^[19]。结合水相单体类型分析可以看出,较低的单体活性需要更长的单体聚合时间,且低活性单体制备的纳滤膜具有低截留特性。高活性单体可快速反应形成致密的有机高分子聚合物,实现纳滤膜对盐离子的有效截留。水相单体的子结构特征 2633 与 2641 对纳滤膜截留性能均有反向作用,降低相应的子结构特征有助于提升纳滤膜截留性能。以上结果表明,通过 SHAP 值法可解析膜制备参数与膜分离效能间的关联,为高性能膜材料开发提供理论方向。

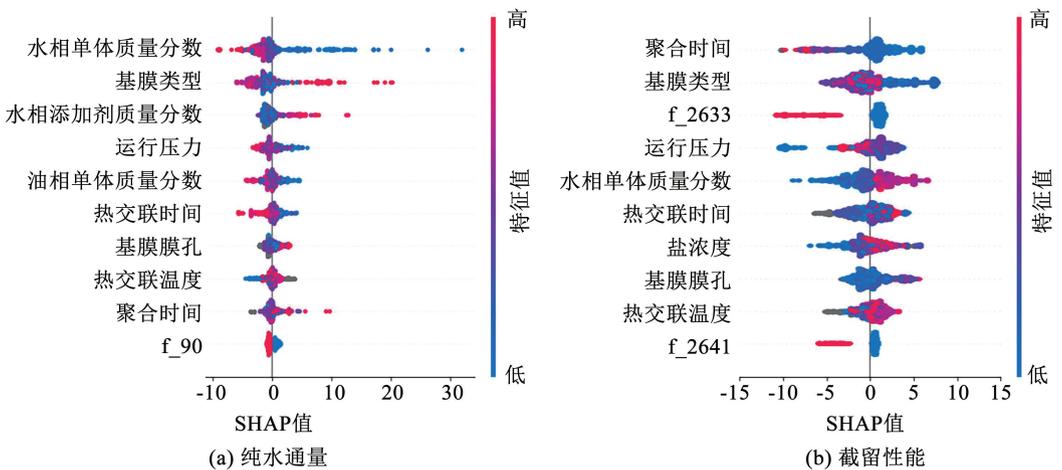


图 4 膜性能 SHAP 图

Fig. 4 SHAP analysis of membrane performance

选取水相单体质量分数、水相添加剂质量分数、油相单体质量分数、聚合时间作为纳滤膜制备的关

键参数,图 5 为 4 种关键参数对膜性能的影响。图 5(a)表明,当水相单体质量分数小于 2% 时,有

助于提升纳滤膜通量;当水相单体质量分数大于2%时,SHAP值趋向于负值,对水通量的促进作用转变为抑制作用。由于水相单体种类多样,难以通过水相单体质量分数判断纳滤膜截留性能。

水相添加剂质量分数分析(图5(b))表明,水相添加剂质量分数的上升有助于纳滤膜纯水通量的提高,以上结果与大部分文献的报道一致。水相添加剂可增大纳滤膜孔径、表面粗糙度、界面亲水性,从而加快水分子的传输^[20]。然而上述特性将在一定程度上降低纳滤膜的截留性能,水相添加剂质量分数的截盐率SHAP值趋向负值。部分数据表明,合理调控水相添加剂质量分数可同步提升纯水通量

与截盐性能。

油相单体质量分数分析(图5(c))表明,降低油相单体质量分数有助于提升纯水通量,然而较低油相单体质量分数不利于盐离子的截留。可以观察到0.1%~0.3%条件下可同时促进纳滤膜纯水通量与截盐率。界面聚合时间(图5(d))分析表明,120 s内时,聚合时间的延长对纯水通量起负面作用,对截留率起促进作用。对于需要更长聚合时间的条件,低活性单体适用于制备高纯水通量与低截盐率的纳滤膜,高活性单体则适用于制备高截盐率的纳滤膜。因此,通过膜制备关键参数的分析,可以基于实际工艺需求设计面向不同应用领域的纳滤膜。

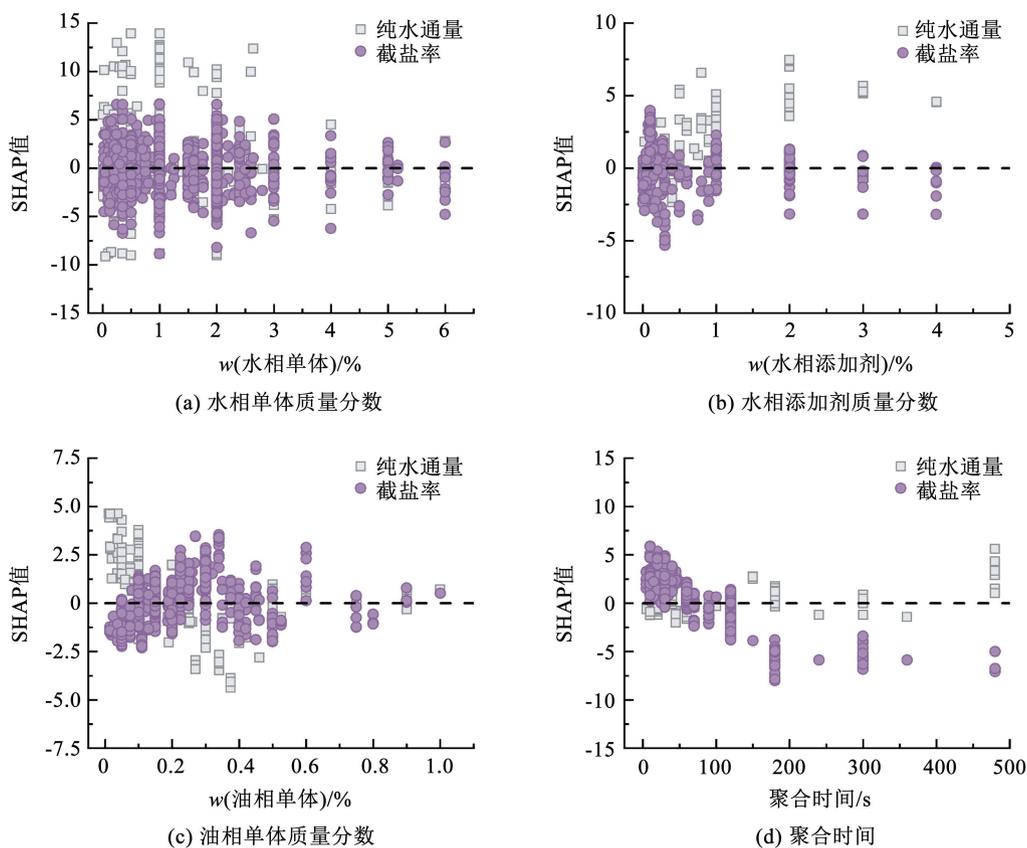


图5 膜制备关键参数对膜性能的影响

Fig. 5 Influence of membrane fabrication parameters on membrane performance

2.4 单体子结构特征解析

基于SHAP模型分析,提取影响纳滤膜性能的关键单体子结构特征。由于数据集中面向纯水通量与截留性能的摩根指纹长度设置不同,相同的子结构特征在不同模型下的特征值并不相同。图6(a)为与纳滤膜纯水通量相关的可视化特征分子指纹。结果表明,子结构特征1831与1674与纯水通量呈正向关系,亲水性基团如羧基(特征1831)与羟基(特征1674)有利于提升膜的纯水通量。同时,有机高分子聚合物内支链(特征1与1380)的增加有助

于增大自由体积分数,加快水分子在纳滤膜内的传输。胺基(特征90与147)为界面聚合的主要官能团,胺基基团的增加促进与酰氯基团的交联,导致纳滤膜的水通量下降。图6(b)为与纳滤膜截留性能相关的可视化特征分子指纹,与纯水通量的结果相反,亲水性基团如羟基(特征2641)对膜的截留性能有负面影响,胺基(特征2741)则促进纳滤膜的截留性能。基于上述分析,推测使用聚乙烯胺、1,2-丙二胺、1-(2-氨基乙基)哌啶等单体有助于开发突破渗透性/选择性制约的纳滤膜。

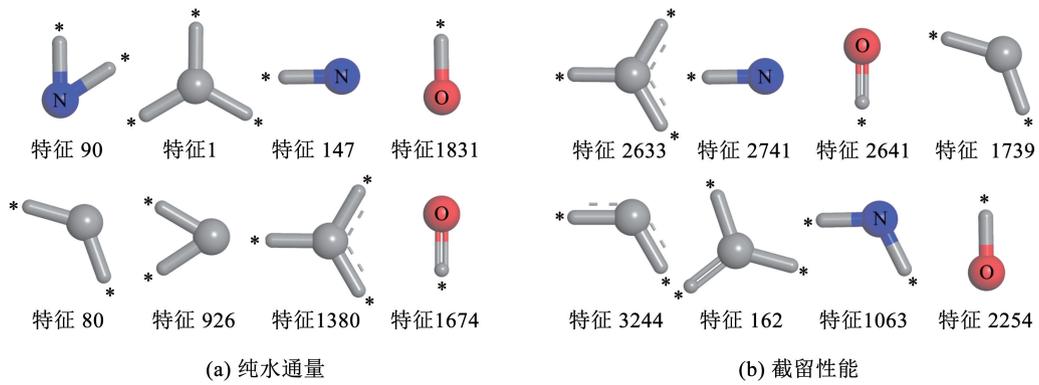


图 6 特征分子指纹

Fig.6 Characteristic atomic groups

3 结 论

1) 基于 XGBoost 与 LightGBM 机器学习模型建立纳滤膜制备参数与纳滤膜性能的关联, 测试集预测能力分析表明, XGBoost 与 LightGBM 机器学习模型对纯水通量的 R^2 评价指标均为 0.84, 对截留性能的 R^2 评价指标分别为 0.90 与 0.86, 表明机器学习模型可基于纳滤膜制备参数有效预测纳滤膜性能。

2) 采用 SHAP 值法对 XGBoost 机器学习模型中的输入参数进行量化, 结果表明, 水相单体质量分数增加对纳滤膜纯水通量有负面影响; 水相添加剂质量分数增加有助于纳滤膜纯水通量的提升, 但对截留性能有反向作用; 降低油相单体质量分数有助于提升纯水通量, 然而较低的油相单体质量分数不利于盐离子的截留。

3) 纳滤膜制备过程中的单体子结构特征分析结果显示, 亲水性子结构特征、支链型子结构特征有助于提升纳滤膜纯水通量; 羟基官能团对膜的截留性能有负面影响, 胺基官能团则促进纳滤膜的截留性能。

参 考 文 献

[1] 梁恒, 李圭白. 饮用水净化工艺的代际认知与融合[J]. 给水排水, 2021, 57(1): 1
LIANG Heng, LI Guibai. Generations' implication and integration of drinking water purification processes [J]. Water & Wastewater Engineering, 2021, 57(1): 1. DOI: 10.13789/j.cnki.wwe1964.2021.01.001

[2] LIN Dachao, BAI Langming, XU Daliang, et al. Effects of oxidation on humic-acid-enhanced gypsum scaling in different nanofiltration phases: performance, mechanisms and prediction by differential log-transformed absorbance spectroscopy [J]. Water Research, 2021, 195: 116989. DOI: 10.1016/j.watres.2021.116989

[3] XU Daliang, ZHENG Junfeng, ZHANG Xin, et al. Mechanistic insights of a thermoresponsive interface for fouling control of thin-film composite nanofiltration membranes [J]. Environmental Science &

Technology, 2022, 56(3): 1927. DOI: 10.1021/acs.est.1c06156

[4] 丁晶, 关淑妍, 赵庆良, 等. 垃圾渗滤液膜浓缩液处理技术研究与应用进展[J]. 哈尔滨工业大学学报, 2021, 53(11): 1
DING Jing, GUAN Shuyan, ZHAO Qingliang, et al. Research and application status of treatment methods of landfill leachate membrane concentrate [J]. Journal of Harbin Institute of Technology, 2021, 53(11): 1. DOI: 10.11918/202104049

[5] XU Daliang, ZHU Xuewu, LUO Xincheng, et al. MXene nanosheet templated nanofiltration membranes toward ultrahigh water transport [J]. Environmental Science & Technology, 2021, 55(2): 1270. DOI: 10.1021/acs.est.0c06835

[6] GAO Haiping, ZHONG Shifa, ZHANG Wenlong, et al. Revolutionizing membrane design using machine learning-bayesian optimization [J]. Environmental Science & Technology, 2022, 56(4): 2572. DOI: 10.1021/acs.est.1c04373

[7] IGNACZ G, BEKE A K, SZEKELY G. Data-driven future for nanofiltration: escaping linearity [J]. Journal of Membrane Science Letters, 2023 (3): 100040. DOI: 10.1016/j.memlet.2023.100040

[8] FLEISCHMAN R B, FARROW K T. Seismic response of perimeter lateral-system structures with highly flexible diaphragms [J]. Earthquake Spectra, 2002, 18(2): 252. DOI: 10.1193/1.1490547

[9] 王晓明, 汪帆, 赵建领, 等. 基于机器学习的斜拉索装配容差区间反演方法[J]. 哈尔滨工业大学学报, 2023, 55(7): 60
WANG Xiaoming, WANG Fan, ZHAO Jianling, et al. Machine learning-based assembly fault-tolerant interval inversion method for stay cables [J]. Journal of Harbin Institute of Technology, 2023, 55(7): 60. DOI: 10.11918/202205118

[10] JEONG N, EPSZTEIN R, WANG Ruoyu, et al. Exploring the knowledge attained by machine learning on ion transport across polyamide membranes using explainable artificial intelligence [J]. Environmental Science & Technology, 2023, 57(46): 17851. DOI: 10.1021/acs.est.2c08384

[11] ZHANG Yanyang, GAO Xiang, SMITH K, et al. Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network [J]. Water Research, 2019, 164: 114888. DOI: 10.1016/j.watres.2019.114888

[12] ZHU Tengyi, ZHANG Yu, TAO Cuicui, et al. Prediction of organic contaminant rejection by nanofiltration and reverse osmosis membranes using interpretable machine learning models [J].

- Science of the Total Environment, 2023, 867: 159348. DOI: 10.1016/j.scitotenv.2022.159348
- [13] GAO Haiping, ZHONG Shifa, DANGAYACH R, et al. Understanding and designing a high-performance ultrafiltration membrane using machine learning [J]. Environmental Science & Technology, 2023, 57 (46): 17831. DOI: 10.1021/acs.est.2c05404
- [14] LU Dan, MA Xuanchao, LU Jiancong, et al. Ensemble machine learning reveals key structural and operational features governing ion selectivity of polyamide nanofiltration membranes [J]. Desalination, 2023, 564: 116748. DOI: 10.1016/j.desal.2023.116748
- [15] 程婉清, 袁定波, 熊鹏, 等. 基于多种机器学习算法的水质指数预测模型构建与评估[J]. 环境科学学报, 2023, 43(11): 144
CHENG Wanqing, YUAN Dingbo, XIONG Peng, et al. Construction and evaluation of city water quality index prediction model based on multiple machine learning algorithms [J]. Acta Scientiae Circumstantiae, 2023, 43 (11): 144. DOI: 10.13671/j.hjkxb.2023.0182
- [16] KARAN S, JIANG Zhiwei, LIVINGSTON A G. Sub-10 nm polyamide nanofilms with ultrafast solvent transport for molecular separation [J]. Science, 2015, 348: 1347. DOI: 10.1126/science.aaa5058
- [17] YIN Jun, YANG Zhe, TANG C Y, et al. Probing the contributions of interior and exterior channels of nanofillers toward the enhanced separation performance of a thin-film nanocomposite reverse osmosis membrane [J]. Environmental Science & Technology Letter, 2020, 7(10): 766. DOI: 10.1021/acs.estlett.0c00507
- [18] 张霄, 董毅, 林赛赛, 等. 机器学习加速能源环境催化材料的创新研究[J]. 能源环境保护, 2023, 37(3): 1
ZHANG Xiao, DONG Yi, LIN Saisai, et al. Machine learning accelerating innovative researches on energy and environmental catalysts [J]. Energy Environmental Protection, 2023, 37(3): 1. DOI: 10.20078/j.eep.20230511
- [19] ZHAO Rui, LI Yi, MAO Yafei, et al. Recycling the high-salinity textile wastewater by quercetin-based nanofiltration membranes with minimal water and energy consumption [J]. Environmental Science & Technology, 2022, 56 (24): 17998. DOI: 10.1021/acs.est.2c06397
- [20] SHAO Senlin, ZENG Fanxi, LONG Li, et al. Nanofiltration membranes with crumpled polyamide films: a critical review on mechanisms, performances, and environmental applications [J]. Environmental Science & Technology, 2022, 56 (18): 12811. DOI: 10.1021/acs.est.2c04736

(编辑 刘 彤)

封面图片说明

封面图片来自本期论文“基于机器学习的纳滤膜预测筛选模型构建与评估”,是基于机器学习对纳滤膜关键制备参数的研究。纳滤净水技术是应对水资源危机和水质安全保障的核心技术之一。然而,纳滤膜性能长期受渗透性与选择性制约,亟需开发高性能纳滤膜。纳滤膜制备过程涉及水相单体质量分数、水相添加剂质量分数、油相单体质量分数、聚合时间等因素,传统的试误实验法需消耗大量的人力、物力与财力。前期研究表明,数据科学与材料研究相结合的机器学习方法具有加速膜材料研发的潜力,利用机器学习可建立膜制备参数-膜分离效能间的关联。通过膜制备过程参数归集,构建基于梯度提升集成学习的纳滤膜预测筛选模型,使用 SHAP 值法解析影响纳滤膜性能的关键参数,分析关键参数对纳滤膜性能的影响机制,明确适用于纳滤膜材料的子结构特征。构建的机器学习模型有助于识别与优化纳滤膜制备过程的关键参数,为纳滤膜的开发提供理论与技术指导。

(图文提供:徐达梁,徐杭镛,靳心瑶,刘超,费兆轩,姚杰,张子峰,李圭白,梁恒. 哈尔滨工业大学环境学院)