

刘项杰,袁巧霞,曹红亮.高湿高脂废弃物产富氢水热油过程的智能建模与响应行为解析[J].华中农业大学学报,2025,44(5):270-279.
DOI:10.13300/j.cnki.hnlkxb.2025.05.027

高湿高脂废弃物产富氢水热油过程的 智能建模与响应行为解析

刘项杰,袁巧霞,曹红亮

华中农业大学工学院,武汉 430070

摘要 为构建水热生物油氢含量高预测精度、强泛化能力的预测模型,深入探讨生物质水热转化规律和机制,以文献中收集的243组病死畜禽、藻类等高湿高脂废弃物水热制备富氢生物油试验数据为基础,采用随机森林和极端梯度提升树2类高适配性机器学习算法,建立高精宽域的水热生物油氢含量预测模型($R^2 > 0.93$);基于数学预测模型,采用可解释技术(Shapley additive explanations, SHAP)及局部依赖性(partial dependence plot, PDP)分析方法,解析水热成油条件的贡献度、局部依赖性响应行为及其互作耦合规律。结果表明:高湿高脂废弃物中的脂质含量与氢含量是制备富氢生物油的决定性因素,二者对油相中氢的富集贡献度排名位居前二,能显著影响油相中氢的积累;随着原料氢含量的增加,油相中氢含量得到提升,表明富氢原料为制备富氢生物油提供了便利条件,提升效果最高可达4%,而原料高位热值是生物油富氢行为的主要抑制因素,抑制作用高达4%。此外,从特征类别间的互作关系来看,元素信息、工业信息及生物质组分信息间的耦合作用强烈,但原料特性与操作条件间的局部耦合作用较小。

关键词 数据驱动;机器学习;随机森林;极端梯度提升树;生物油;热值

中图分类号 S216.2;X705 **文献标识码** A **文章编号** 1000-2421(2025)05-0270-10

病死畜禽以及藻类等高湿高脂废弃物无害化处理与资源化利用,对于我国畜禽养殖业可持续发展、水环境保护以及美丽乡村建设等具有重要现实意义^[1-3]。水热转换技术以水分为溶解介质,在高温高压条件下将有机物转化为水热炭、水热生物油以及CO、H₂、CO₂、水蒸气等气相组分,其中水热炭、水热油、CO、H₂作为可燃性产物,二次利用可实现有机废弃物的资源化利用^[4-6];且病死畜禽、藻类等高脂废弃物中存在大量油脂类成分,是重要的水热生物油前驱体,利于制备高品质的水热生物油^[7-9]。

水热生物油中氢元素含量是体现燃油品质的一个重要指标,直接影响生物油的燃烧热值、闪点以及密度等品质特性^[10-11]。为提升水热生物油氢含量,国内外相关研究从反应装置设计^[12]、催化剂开发^[13]、水热条件优化^[14]等方面进行了大量探索研究。Wagner等^[15]设计了一种连续型水热反应器来提升藻类生物质制备生物油的产量及品质,结果表

明提高加热速率和反应时间,连续型反应器能显著提升生物油产量,同时在321.6 °C时油相中的氢回收率最高,表明反应器在特定条件下能促进氢元素向油相转化;Shakya等^[16]研究了300、350 °C下5种催化剂对藻类生物油升级效果,结果表明300 °C下获得了高产量的生物油,350 °C下获得了高品质的生物油,二者油相中的氢含量均得到了一定程度的提升;Basar等^[17]通过总结水热液化技术的关键设计和操作参数来优化整体工艺性能,结果表明提升加热速率、降低保温时间是生物油生产的最佳液化工艺,对其进行调控可实现对整体工艺性能的优化。上述结果表明,无论是反应装置结构设计,还是相关催化剂的开发与操作条件优化,均需进行大量的试验来助力于油相品质的提升,较为耗时、耗力。

机器学习技术凭借其强大的自动化决策能力、预测分析能力、大数据处理能力、强大的适应性使其有助于水热液化制备生物油过程与工艺的优化,为

收稿日期:2024-07-04

基金项目:湖北省技术创新计划重点研发专项(2023CA153)

刘项杰,E-mail:654447349@qq.com

通信作者:袁巧霞,E-mail:qxyuan@mail.hzau.edu.cn

此部分学者进行了机器学习辅助水热液化技术的研究。Zhang等^[18]通过机器学习技术实现了生物油生产的预测和优化,结果表明,操作条件(温度和停留时间)的相对重要性高于藻类特性。Katongtung等^[19]利用机器学习方法实现了湿生物质水热液化得到的生物油产量和高位热值的预测,研究发现温度是对产量和高位热值最有影响力的特征参数,原料特性对模型的贡献超过55%,该研究揭示了最重要特征对预测目标的影响和相互作用,从而更好地理解了水热液化反应系统。但现有文献中将机器学习结合到制备富氢生物油燃料的研究较少,因此建立一个能实现不同工艺结构及操作条件下水热生物油氢含量准确预测的数学模型,对高品质生物油的生产具有重要的指导意义,尤其是亟需预测精度高、泛化能力强的预测模型^[20-21]。

本研究在课题组前期研究积累基础上,同时调研国内外相关研究资料,收集并建立1个病死畜禽、藻类等高湿高脂废弃物水热成油的数据集;同时,通过选用具有高适配性的随机森林和极端梯度提升树2类机器学习算法,结合MissForest数据填充、5折交叉验证等建模手段,获得1个高精宽域的水热生物油氢含量预测模型。基于高精宽域数学预测模型,采用SHAP可解释技术及局部依赖性分析方法,深入

解析水热成油条件的贡献度、局部依赖性响应行为及其互作耦合规律,旨在为高湿高脂废弃物制备富氢水热生物油提供重要理论支撑。

1 材料与方法

1.1 数据集及预处理方法

数据集主要由病死畜禽、藻类等高湿高脂废弃物原料制备水热生物油的实验数据构成,共计243组(其中病死畜禽原料58组,藻类原料185组);数据集包含原料的元素信息、生物质组分信息、生物质工业分析成分信息以及生物油氢含量等19个特征指标。

利用Python对数据集进行统计分析,针对数据集中部分数据缺失问题,采用MissForest算法进行数据填充处理;为探索不同输入特征对模型预测精度的影响,根据不同的特征类别重新整合填充,构建4个新数据集(表1);训练数据集与测试数据集按照9:1随机划分,确保建模数据的随机性;为降低数据量纲的不一致性、数据分布的不均匀性对建模精度的影响,对数据集进行Z score标准化变换:

$$X_i^* = \frac{X_i - \mu}{\sigma} \quad (1)$$

其中, X_i^* 是每个输入变量的标准化值, X_i 是输入变量的原始实验数据值, μ 和 σ 分别是输入特征 X_i 的平均值和标准差。

表1 生物油氢含量预测模型对应数据集特征情况

Table 1 Characterization of the corresponding dataset for the bio-oil hydrogen content prediction model

数据集 Dataset	模型 Model	特征数 Characteristic number	输入 Inputs	目标 Target
B_H_Data1	B_H_RF1	16	EC、PC、OC	B_H
B_H_Data2	B_H_RF2	15	EC、BC、OC	B_H
B_H_Data3	B_H_RF3	13	PC、BC、OC	B_H
B_H_Data4	B_H_RF4	19	EC、PC、BC、OC	B_H
B_H_Data1	B_H_XGBoost1	16	EC、PC、OC	B_H
B_H_Data2	B_H_XGBoost2	15	EC、BC、OC	B_H
B_H_Data3	B_H_XGBoost3	13	PC、BC、OC	B_H
B_H_Data4	B_H_XGBoost4	19	EC、PC、BC、OC	B_H

注 Note: B_H_Data:生物油氢含量数据集 Bio-oil hydrogen content dataset; B_H:生物油含氢量简称 Abbreviation for bio-oil hydrogen content; EC、PC、BC、OC 分别为元素组成、工业组成、生物质组成、操作条件的英文名称简写,生物质组成包括脂质含量、蛋白质含量、碳水化合物含量 Abbreviations for the English names of elemental composition, proximate composition, biomass composition, and operating conditions, respectively. Biomass composition: lipid content, protein content, and carbohydrate content.

1.2 模型开发及评估方法

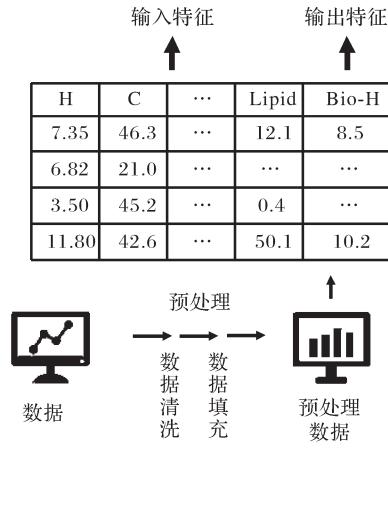
数据预处理后,结合数据的特性,采用2种经典的集成学习算法随机森林(RF)和极端梯度提升树(XGBoost)构建富氢生物油的数学预测模型;建模方法及流程如图1所示,包括数据预处理、数据集划分、

超参数优化、模型验证等关键步骤。

为提高模型整体的稳健性,在训练集数据上采用5折交叉验证技术结合贝叶斯优化技术实现2种算法模型的超参数优化,选取最优超参数作为建模算法的输入参数,构建最终生物油氢含量预测模型;

除随机森林(树的数量 n_estimators、最大树深度 max_depth)和 XGBoost(学习率 eta、最大树深度 max_depth)等确定的超参数外,随机森林算法和XGBoost算法的其他超参数均保持默认值。随后,利用测试集对模型的效果进行验证。

模型性能评估采用回归系数(R^2)和均方根误差(RMSE,公式中用 R_{MSE} 表示)进行通用性评价:



$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y_i^{\text{pre}})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (2)$$

$$R_{\text{MSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y_i^{\text{pre}})^2} \quad (3)$$

式(2)~(3)中, N 表示样本数, y_i , y_i^{pre} 和 \bar{y} 分别表示实际数据、估计值和平均值。 R^2 值表示总体数据的拟合程度。

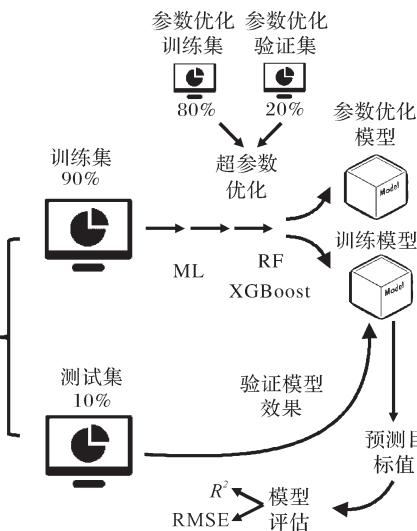


图1 模型开发流程

Fig.1 Flow of bio-oil hydrogen content prediction model development

1.3 特征重要度评估与解释方法

在建立稳定性高且精准的预测模型基础上,采用可解释技术(Shapley additive explanations, SHAP)和局部依赖性(partial dependence plot, PDP)分析方法提取特征及预测目标间的关联性,对原料特性和水热液化条件等信息的相对重要性进行排序,评估特征信息在实验指导方面的有效性。并结合特征的边际影响力,展现特征对目标的影响趋势,从而实现对黑盒模型的结果进行解释。

2 结果与分析

2.1 数据集统计学分布规律及相关性

对收集到的数据进行统计分析,其中EC、PC、BC部分信息分别展现了原料的元素、工业组分及生物质组成信息等原料特性数据的统计结果。OC、B_H分别表示试验操作条件及生物油氢含量信息等试验特性数据的统计结果。

为更深入了解数据情况,利用数据分布小提琴图及相关性热力图(图2)探究数据深层信息。由图2A可知,碳、氧元素虽占据元素组成的主导地位,但

碳元素的主体部分集中在46.30%,而氧元素的主体集中在27.90%。除氧元素外,其余元素组成部分数据均表现出一定程度上的波动性,氮元素的波动性尤其明显,其数据分布存在多个峰谷点,但其数据均匀度也属元素数据中最高,可见数据收集时具有较好的均匀性。

水分和挥发成分的主体部分分别出现在73.20%和61.80%,而从整体来看工业组分信息均存在较大的数据波动性,尤其是原料高位热值波动极其明显,表明高位热值受到多因素影响,数据稳定性较差(图2B)。图2C展示试验操作条件的分布趋势,可知反应时间及反应压力的波动性较为明显,反应温度、加热速率、反应时间、反应压力及固液比分别主要分布在300.00 °C、15.00 °C/min、60.00 min、12.00 MPa、14.00%左右。

图2D展示了生物质组分中脂肪、蛋白质、碳水化合物的主体分布位置,分别在12.05%、37.00%及22.90%,三者数据整体分布具有波动性,蛋白质组分尤其明显,且与氮的分布趋势非常相似,推测二者间存在关联性导致数据分布呈现出较高的一致性。总

体来看,原料特性及操作条件数据分布均展现出数据收集的广度和丰度,也佐证了原料高含水特性的存在,但大部分数据的分布存在波动性及相似性,表明数据因素间存在一定程度的关联性。而图2E也证实油中氢含量与原料中的碳、氢和脂质含量表现出明显的正相关性($r \geq 0.54$),而与灰分、挥发分和保温时间存在显著的负相关性($r \leq -0.45$),进一步证实

了因素与目标之间的关联。同时关联性的复杂性与大部分信息的非线性关联性,也表明生物油氢含量与原料特性及操作条件间的复杂关系。

数据展现的分布规律及因素间的关联性表明整个水热反应系统是一个复杂的非线性系统,因此,选取非线性关联数据适配性优越的树形结构的集成学习算法(随机森林、XGBoost)构建模型。

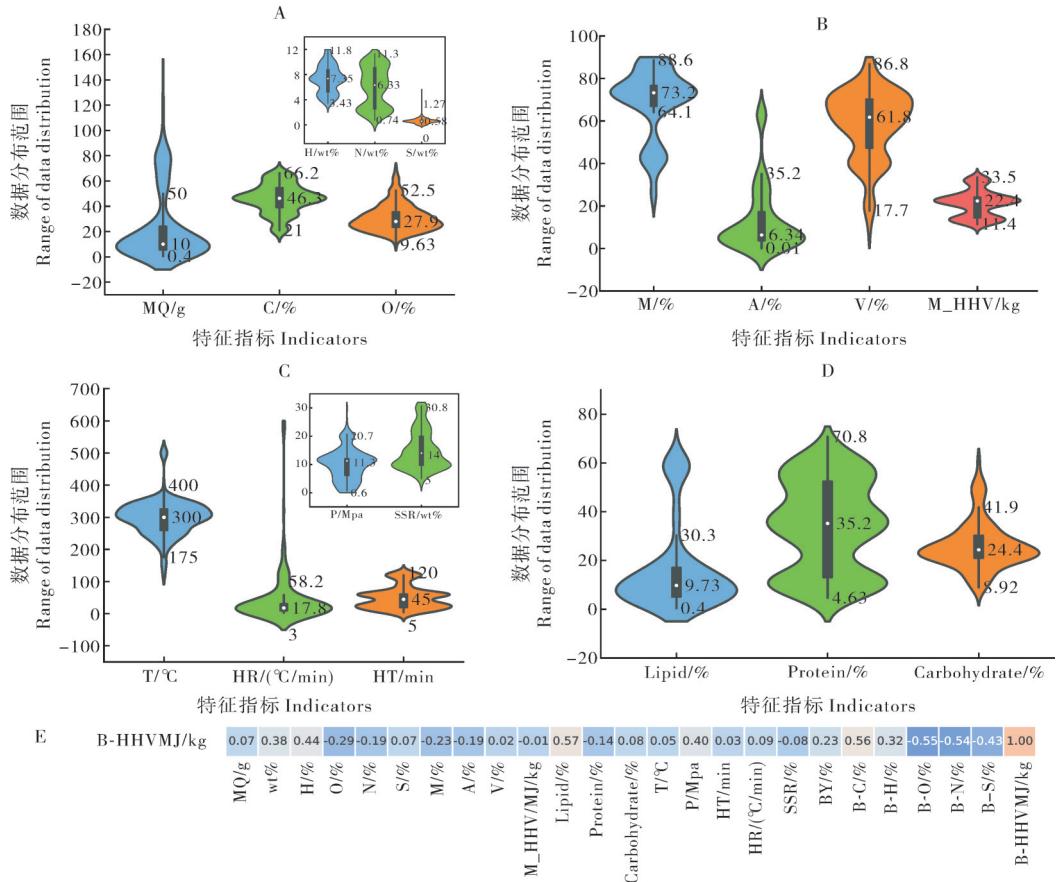


图2 数据集分布规律及相关性图

Fig.2 Distribution pattern and correlation diagram of the dataset

2.2 机器学习模型性能评价

构建的生物油中氢含量模型性能如表2所示。XGBoost 和 RF 算法建立的模型都取得了较为优异的结果(R^2 : 0.932~0.998; RMSE: 0.005~0.498)。

对比富氢生物油预测模型的测试集回归系数(R^2)可知,B_H_XGBoost1相比其他所有模型均表现出更强建模预测能力。较高的 R^2 表明元素和工业组分信息与水热液化技术参数的结合能有效促进模型效果。不仅如此,其建立的模型均方根误差亦最低,充分展现了模型泛化性能及对误差的控制能力。从数据角度对比测试模型可知,B_H_Data1对应模型表现出最佳拟合效果(R^2 : 0.947, 0.960),而B_H_Data2对应模型效果最差(R^2 : 0.932, 0.937)。

表2 生物油氢含量预测模型性能

Table 2 Bio-oil hydrogen content prediction model performance

Model 模型	训练 Train		测试 Test	
	R^2	RMSE	R^2	RMSE
B_H_RF1	0.970	0.303	0.947	0.416
B_H_RF2	0.985	0.236	0.932	0.498
B_H_RF3	0.983	0.251	0.944	0.467
B_H_RF4	0.982	0.245	0.936	0.472
B_H_XGBoost1	0.974	0.285	0.960	0.358
B_H_XGBoost2	0.998	0.089	0.937	0.498
B_H_XGBoost3	0.996	0.126	0.947	0.475
B_H_XGBoost4	0.995	0.005	0.948	0.446

B_H_Data3、B_H_Data4 对应模型均表现中等影响 (R^2 : 0.944, 0.947; 0.936, 0.948)。B_H_Data2 将生物质成分特征替换 B_H_Data1 中的近似分析特征, 导致模型性能下降, 表明生物质成分比工业组分对模型的影响力弱。B_H_Data3 用近似成分数据替换了 B_H_Data2 中的元素成分数据, 导致模型效率提升, 证明近似成分数据的有效性高于元素数据。对比 B_H_Data1 和 B_H_Data3, 可以得出结论, 元素信息比生物量成分信息为模型提供了更有利的提升。因此, 总体结果表明, 除了特征因子中水热液化技术操作参数外, 所有 3 组原料特征都支持生物油产量预测模型。这些组分对模型效果影响力高低排序为: 近似成分信息 > 元素信息 > 生物质信息数据。同时由于上述分析时, 模型差异极小, 为确保推测的准确性, 还需结合训练集模型的效果来进一步证明推断。

从训练集对应结果来看 B_H_Data2 建立的模型在 2 种算法模型上均取得最为优异效果。模型训练集结果表明, B_H_Data1 的效果最差 (R^2 : 0.970, 0.974), B_H_Data3、B_H_Data4 对所有模型影响均衡 (R^2 : 0.983, 0.982)。在此基础上对比不同数据集训练模型结果可知, B_H_Data2 将生物质成分特征替换 B_H_Data1 中的近似分析特征, 训练模型的效果得以提升, 表明生物质成分比工业组分对模型的影响更显著。B_H_Data3 用近似成分数据替换了 B_H_Data2 中的元素成分数据, 模型效率下降, 这表明近似成分数据的有效性低于元素数据。比较 B_H_Data1、B_H_Data3 可知, 元素信息比生物质成分信息为模型提供了更不利的影响, 因此, 可知 3 种信息组分对训练模型效果的影响力排序依次为生物质组分 > 元素组成 > 工业组成信息。

综合训练集和测试集中特征类别对模型效果的影响力排序可知, 元素信息在整个建模过程中的重要性, 而工业组分信息及生物质信息组分对模型的影响过程更为复杂, 会随着模型算法及数据集变化, 并使得最终模型产生不同效果。故选择最优模型时, 需综合考虑训练集和测试集模型的效果。由于训练模型的 R^2 能够代表模型学到的真实信息效果, 其值越高代表模型学到的规律更为真实且可信。而测试集 R^2 表达了模型对新规律的接受能力与包容性, 结合元素信息的重要性以及工业组分信息的复杂影响, 最终选取训练集、测试集效果优异, 且对特征信息包容度更强的 B_H_XGBoost4 模型作为最优模型, 为后续进一步探究特征间的反应机制提供基础。

依据富氢生物油预测模型效果的评估及分析, 最终选取最优预测模型为 B_H_XGBoost4, 但未能推断出 3 种组分对模型推进力大小的排序, 仅发现元素组成对模型的极高重要度。

2.3 基于预测模型的水热条件贡献度分析

最优预测模型的评估, 表明元素信息的重要性, 其余信息的影响力不清, 故结合特征贡献度进一步探究特征因素对目标特性的影响。富氢生物油预测模型的特征重要度排序及特征值影响力散点分布如图 3 所示, 在全局因素中, 脂质特性和原料中氢含量具有显著影响力, 而原料的高位热值及碳、蛋白质、灰分、氧元素和挥发分含量对生物油中富氢行为的影响力依次降低。此外, 其余特征仍会对生物油氢含量的富集造成影响, 但其影响力较弱。总体来看, 3 个层次中因素分布表明影响模型效果的因素包含元素组成、工业组成、生物组成等分类特征中 2 种以

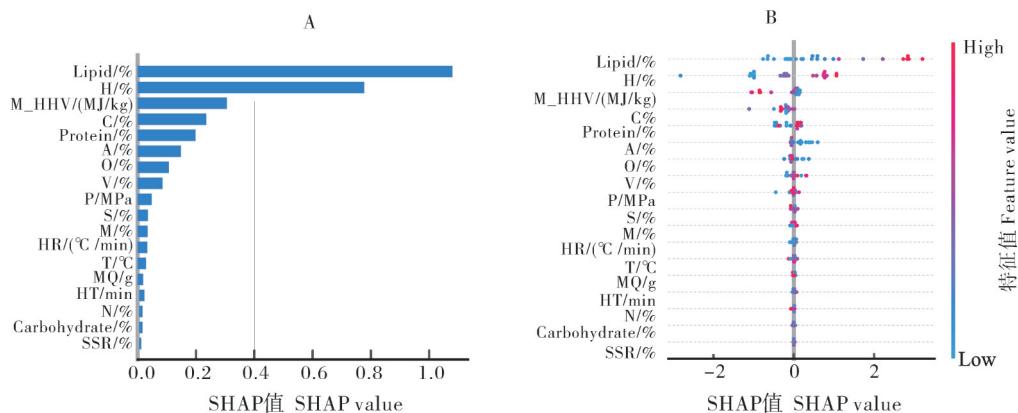


图 3 生物油氢含量最优模型 SHAP 特征重要度直方图及散点响应图

Fig.3 Histogram of importance of SHAP features and scatter response plots of the optimal model for hydrogen content of bio-oils

上的因素特征。由此可见特征对油中氢含量的影响是极其复杂的,也直观地印证了建模表观分析得出的推断。

影响力排序直方图(图3A)展现了各因素对生物油氢含量的相对影响力大小,但其对目标的推进方向不清晰。而特征重要性散点图(图3B)能清晰且直观地解决影响力排序直方图的局限性。通过特征重要性散点图,可知高含量的脂质和氢元素特征对生物油氢含量具有促进作用,低含量的脂质和氢含量会抑制生物油中氢的积累。由此可见富含脂质和氢的原料在制备富氢生物油的过程具有显著优势。相对于重要度前2个特征量对高含量的促进作用,中间6个特征量对生物油中氢含量未表现出明显的的影响。且其中蛋白质、挥发分的含量变化对生物油氢含量的影响趋势混乱,推断是由于部分因素间存在交互影响,导致因素对目标的影响不稳定。故后续需要关注因素间交互作用对目标特性的影响。后10个特征量贡献度排序表明其对生物油氢含量影响力微弱,特征含量的变化,未对目标提升起到明显

作用。

SHAP特征重要性排序直方图及重要性散点图展现了特征因素对目标影响力排序及方向,但其仍停留在定性分析程度上,未能精确分析具体特征变化对目标值的影响。故为降低模型算力及时间成本,后续将围绕前8个特征量,着重从定量分析角度对试验流程进行更为精确的研究,探究试验过程的响应行为。

2.4 水热条件单因素局部依赖性响应行为

前8个特征量对油中氢含量局部依赖性结果(图4)表明,脂质含量和原料氢含量对生物油氢含量提升展现出显著潜力,随着其含量提升,最终产物生物油中的氢含量也会不断提高。且从潜力极限来看,二者均能使得目标值提升4%,但脂质含量对目标值的提升更为迅速,且整体趋势较为平稳。相对于脂质含量和原料氢含量的提升潜力,后3个特征量对目标值的极限影响力除原料高位热值外均低于1%,且除原料高位热值、灰分、氧与目标值的响应外,整体变化趋势并不平滑,表现出一定程度的波动性。

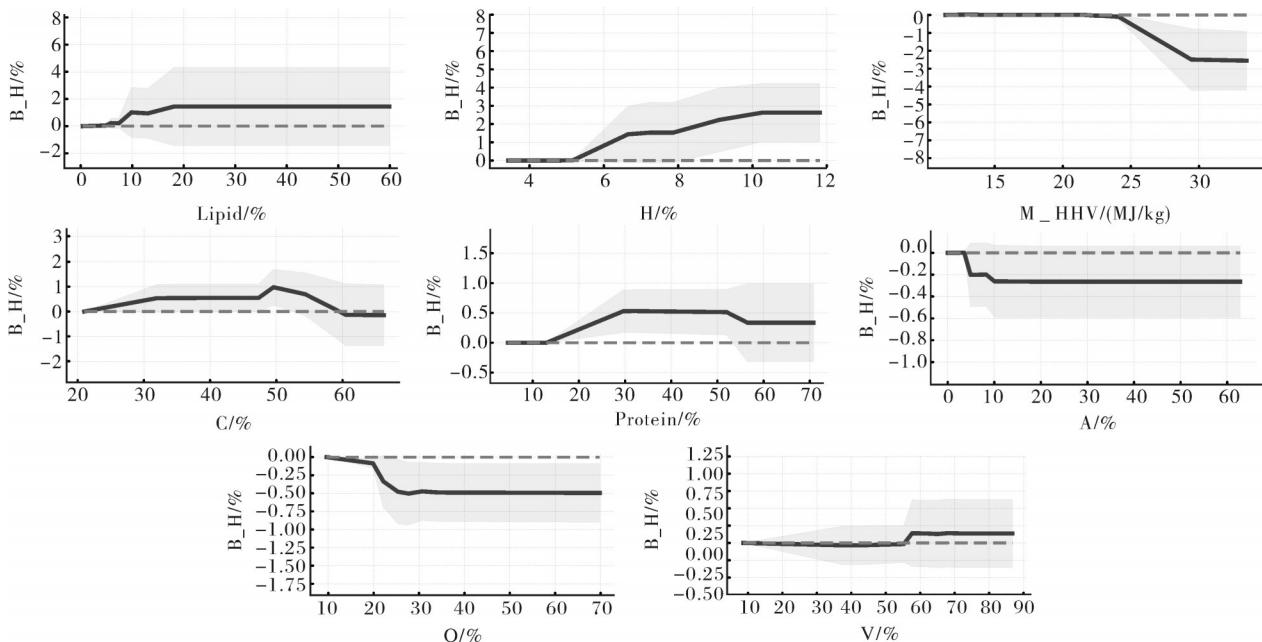


图4 单因素局部依赖性趋势图

Fig.4 Trends in single-factor localized dependencies

整体观察两梯队特征影响趋势可知,仅原料高位热值、灰分和氧含量对生物油氢含量的影响是负方向的,其余特征均表现一定强度推进作用。原料高位热值在25 MJ/kg级别以上对生物油氢含量展现出强抑制性,随着热值提升,抑制效果最高能达到4%,推测是由于高热值原料丰富的挥发性物质,随

着反应进行,通过脱水作用形成含氢类气体^[22],降低含氢有机物总量,最终使得进入油相中氢含量降低^[23]。而灰分(0%~10%)和原料氧含量(10%~30%)的提升均导致目标值有少许降低,但在此之后对目标的影响趋于稳定,这主要是由于灰分和氧含量提升,意味着整体原料中的稳定物质增加、有机物

降低,使得可反应物原料含量降低^[24],同时有研究表明水热转化过程中的氧元素主要迁移到油相以外的三相产物中^[22],抑制了原料有机物中氢元素向油相中转化的反应过程,最终使得油相中氢含量降低,而在极高值时的平稳性可能是反应物质消耗完全,反应达到稳态。反观其余正向推动反应特征可知,除挥发分外其余特征均能表现出原料中有机物含量的提升,促进了反应正向进行,进而使得油相中氢含量提升。而挥发分含量的增长在55%以前并未表现出明显推进效果,在55%~60%范围内展现出明显促进作用,推断原因为原料中挥发性物质增多,提升了挥发物中含氢化合物向油相转化进程,使得油相氢含量得到提升^[25]。

总体而言,单因素局部依赖性特征分析清晰表明了因素与目标值间的响应趋势,其中脂质含量和原料中氢元素含量能有效提升生物油中的氢含量,其提升效率能达到4%;而原料高位热值对生物油氢

含量抑制作用最高,最高可达4%,平均可达2.8%,可见其为抑制油相氢含量的重要因素。同时全局因素间趋势中部分因素展现的波动性表明,趋势变化中存在干扰信号,后续需进一步探究因素间交互作用机制。

2.5 水热条件间互作耦合规律

由单因素特征及目标响应行为可知除原料高位热值、灰分和氧含量外其余特征均展现出趋势的波动性,可认为各因素间存在相互影响的耦合机制。故优先选取重要度强及波动性极其明显的因素进一步探究其间存在的耦合关系(氢和蛋白质)及其对生物油氢含量值的影响效应大小。

氢含量与其余特征间交互作用如图5A所示,大部分特征间均未展现出显著耦合作用,只有碳、原料高位热值、脂质、蛋白质及压力与氢之间均存在较为显著的耦合作用。其中又以碳和原料高位热值作用效果尤为显著。单看氢、碳对生物油氢含量的影响,

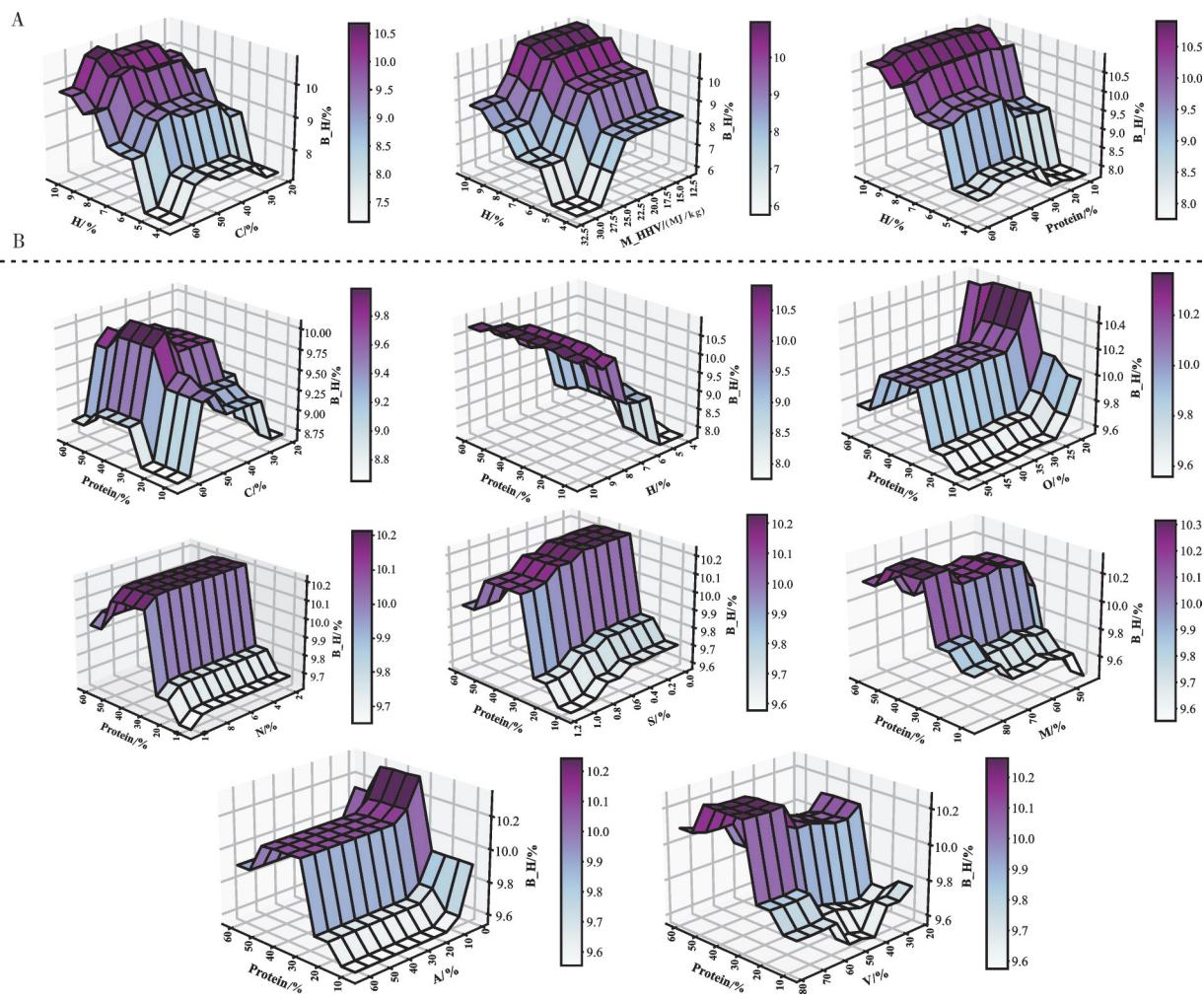


图5 氢元素、蛋白质显著交互因素的影响趋势图

Fig.5 Trend plot of the effect of significant interaction factors of elemental hydrogen, protein

可知二者间存在明显的耦合作用,且当碳在中等含量(50%~60%),氢在高含量(10%)时,生物油氢含量取得最高值(10.05%左右)。由此可见,较高的碳含量使得参加反应的物质原料得到提升,能一定程度促进反应进行,但其影响力显著小于原料中氢含量对目标值的作用,推断原因可能在于相对于含氢元素的物质,含碳元素物质存在更为稳定的结构,使得对目标的助力大小受限,降低了其影响力。而原料高位热值和氢元素的交互作用表明,在低热值原料(12.5%~22.5%)和高氢含量原料的前提下,生物油氢含量达到10%左右的最高值。低热值原料决定了生物油氢含量的极值,而原料氢含量高低决定了目标值的最值。除上述2个特征外,脂质、蛋白质及压力虽表现出一定程度的耦合关联,但其均出现在特征较低值的前提下,随着氢含量不断提升,趋势逐渐被拉回到氢的影响水平上。全局趋势表明,低氢含量原料不能主导反应的进行方向,而随着原料富氢量的提升,氢元素主导作用逐渐显现,表明富氢原料有利于富氢生物油的生产。

不同于氢和特征间的交互响应,蛋白质与其余特征间的交互作用如图5B所示。除原料高位热值、生物质组分信息及操作条件外,与其他因素间均展现出明显耦合作用,其中与碳和挥发分又展现出复杂的变化趋势。蛋白质和碳对生物油氢含量的影响表明,油相中氢含量最高值(10%)出现在中等含量蛋白质(30%~50%)及碳含量(50%)的条件下,且碳比蛋白质展现出更强的影响力,原因在于碳含量与蛋白质含量呈现包容关系,间接决定了原料有机物的含量,使得原料中蛋白质含量相对增加,提高了酯化反应的强度^[26],加速了物质向油相转化,油相中氢含量得到提升。蛋白质的影响曲线表明仅中等量级的蛋白质含量能促进目标提升,推测主要是由于蛋白质种类复杂,大多数高分子质量的蛋白质种类由于肽键等稳定结构^[27],转化时需要获得更大的活化能,故高含量的蛋白质使得目标值不升反降。相较于碳耦合作用趋势的凸起性,挥发分的耦合作用趋势呈现一定程度凹陷性,尤其是当挥发分含量处于40%~60%范围内,表现出显著凹陷性,推断主要原因是蛋白质在此段区域内含量很低,而挥发性物质的增多使得反应析出更多的碳氢化合物等挥发性成分^[28],油相中的氢元素降低,随着蛋白质含量增加,向油相转化的方向重新获取主导权^[29],进而油相中氢含量得到一定程度提升。

总体观察生物油氢含量与氢、蛋白质及其余特征间耦合作用趋势图,可知耦合作用发生的情况大多出现在原料特性间,原料特性与操作条件间的局部耦合作用发生概率较低。故对整体工艺优化时,需要重点关注原料的氢和脂质以及挥发分含量。

3 讨 论

本研究通过对比研究具有高适配性的随机森林和极端梯度提升树两类机器学习建模算法,搭建了高精度、强泛化能力的高湿高脂农业废弃物产富氢水热燃油过程的智能预测模型($R^2>0.948$)。从水热条件特征贡献度来看,全局因素中原料脂质含量与氢含量是制备富氢生物油的决定性因素,且二者均对油相中氢的富集起到促进作用。单因素响应行为趋势表明,富氢原料为制备富氢生物油提供了便利条件,其对生物油氢含量的提升效率最高可达4%。而原料高位热值是生物油富氢行为的主要抑制因素,抑制效果最高可达4%,平均可达2.8%。双因素耦合作用趋势表明,富氢原料、蛋白质与其余特性间的耦合趋势复杂。而分类别来看,元素信息、工业信息及生物质组分信息间耦合作用频发,原料特性与操作条件间的局部耦合作用发生概率较低。

参考文献 References

- [1] SAMORAJ M, MIRONIUK M, IZYDORCZYK G, et al. The challenges and perspectives for anaerobic digestion of animal waste and fertilizer application of the digestate [J/OL]. Chemosphere, 2022, 295: 133799 [2024-07-04]. <https://doi.org/10.1016/j.chemosphere.2022.133799>.
- [2] 武欣艳,李鑫,于洁,等.水环境中藻类与细菌有益相互作用的研究进展与展望[J].华中农业大学学报,2024,43(4):41-50. WU X Y, LI X, YU J, et al. Progress and prospects on studying beneficial interactions between algae and bacteria in aquatic environments [J]. Journal of Huazhong Agricultural University, 2024, 43(4): 41-50 (in Chinese with English abstract).
- [3] ZHANG X, ZHOU T, YANG L Y, et al. Mechanism of the effect of temperature on the characteristics of microwave hydrothermal products of dead pig model compounds [J/OL]. Journal of analytical and applied pyrolysis, 2024, 179: 106458 [2024-07-04]. <https://doi.org/10.1016/j.jaat.2024.106458>.
- [4] HUSSIN F, HAZANI N N, KHALIL M, et al. Environmental life cycle assessment of biomass conversion using hydrothermal technology: a review [J/OL]. Fuel processing technology, 2023, 246: 107747 [2024-07-04]. <https://doi.org/10.1016/j.fuproc.2023.107747>.
- [5] WANG H, LI H, HAN X, et al. Biomass conversion by hydro-

- thermal liquefaction technology [M]. Amsterdam: Elsevier, 2023.
- [6] SHAHBEIK H, KAZEMI SHARIAT PANAHY H, DEHHAGHI M, et al. Biomass to biofuels using hydrothermal liquefaction: a comprehensive review [J/OL]. Renewable and sustainable energy reviews, 2024, 189: 113976 [2024-07-04]. <https://doi.org/10.1016/j.rser.2023.113976>.
- [7] KIRAN KUMAR P, VIJAYA KRISHNA S, VERMA K, et al. Bio oil production from microalgae *via* hydrothermal liquefaction technology under subcritical water conditions [J]. Journal of microbiological methods, 2018, 153: 108-117.
- [8] ZHU Y H, ALBRECHT K O, ELLIOTT D C, et al. Development of hydrothermal liquefaction and upgrading technologies for lipid-extracted algae conversion to liquid fuels [J]. Algal research, 2013, 2(4): 455-464.
- [9] LIU Q, KONG G, ZHANG G Y, et al. Recent advances in hydrothermal liquefaction of manure wastes into value-added products [J/OL]. Energy conversion and management, 2023, 292: 117392 [2024-07-04]. <https://doi.org/10.1016/j.enconman.2023.117392>.
- [10] MISHRA R K, KUMAR V, KUMAR P, et al. Hydrothermal liquefaction of biomass for bio-crude production: a review on feedstocks, chemical compositions, operating parameters, reaction kinetics, techno-economic study, and life cycle assessment [J/OL]. Fuel, 2022, 316: 123377 [2024-07-04]. <https://doi.org/10.1016/j.fuel.2022.123377>.
- [11] KARTHIC S V, SENTHIL KUMAR M, NATARAJ G, et al. Experimental investigations on the influence of hydrogen and LPG mixtures on performance behavior of a Mahua bio-oil-powered dual fuel engine [J]. International journal of green energy, 2019, 16(12): 878-889.
- [12] MAQBOOL W, BILLER P, ANASTASAKIS K. A kinetic process model for sewage sludge hydrothermal liquefaction in Aspen Plus: model validation with pilot-plant data and scale up [J/OL]. Energy conversion and management, 2024, 302: 118136 [2024-07-04]. <https://doi.org/10.1016/j.enconman.2024.118136>.
- [13] RIBEIRO F C P, SANTOS J L, ARAUJO R O, et al. Sustainable catalysts for esterification: sulfonated carbon spheres from biomass waste using hydrothermal carbonization [J/OL]. Renewable energy, 2024, 220: 119653 [2024-07-04]. <https://doi.org/10.1016/j.renene.2023.119653>.
- [14] TANG Q H, CHEN Y Q, YANG H P, et al. Prediction of bio-oil yield and hydrogen contents based on machine learning method: effect of biomass compositions and pyrolysis conditions [J]. Energy & fuels, 2020, 34(9): 11050-11060.
- [15] WAGNER J L, LE C D, TING V P, et al. Design and operation of an inexpensive, laboratory-scale, continuous hydrothermal liquefaction reactor for the conversion of microalgae produced during wastewater treatment [J]. Fuel processing technology, 2017, 165: 102-111.
- [16] SHAKYA R, ADHIKARI S, MAHADEVAN R, et al. Catalytic upgrading of bio-oil produced from hydrothermal liquefaction of *Nannochloropsis* sp. [J]. Bioresource technology, 2018, 252: 28-36.
- [17] BASAR I A, LIU H, CARRERE H, et al. A review on key design and operational parameters to optimize and develop hydrothermal liquefaction of biomass for biorefinery applications [J]. Green chemistry, 2021, 23(4): 1404-1446.
- [18] ZHANG W J, LI J, LIU T G, et al. Machine learning prediction and optimization of bio-oil production from hydrothermal liquefaction of algae [J/OL]. Bioresource technology, 2021, 342: 126011 [2024-07-04]. <https://doi.org/10.1016/j.biortech.2021.126011>.
- [19] KATONGTUNG T, ONSREE T, TIPPAYAWONG N. Machine learning prediction of biocrude yields and higher heating values from hydrothermal liquefaction of wet biomass and wastes [J/OL]. Bioresource technology, 2022, 344: 126278 [2024-07-04]. <https://doi.org/10.1016/j.biortech.2021.126278>.
- [20] LIU Q, ZHANG G Y, YU J J, et al. Machine learning-aided hydrothermal carbonization of biomass for coal-like hydrochar production: parameters optimization and experimental verification [J/OL]. Bioresource technology, 2024, 393: 130073 [2024-07-04]. <https://doi.org/10.1016/j.biortech.2023.130073>.
- [21] ZHANG W J, CHEN Q Y, CHEN J F, et al. Machine learning for hydrothermal treatment of biomass: a review [J/OL]. Bioresource technology, 2023, 370: 128547 [2024-07-04]. <https://doi.org/10.1016/j.biortech.2022.128547>.
- [22] 朱张兵.农村典型废弃物水热转化效率与产物特性研究[D].北京:中国农业大学,2018.ZHU Z B. Study on hydrothermal conversion efficiency and product characteristics of typical rural wastes[D]. Beijing: China Agricultural University, 2018(in Chinese with English abstract).
- [23] STANGER R, XIE W, WALL T, et al. Dynamic measurement of coal thermal properties and elemental composition of volatile matter during coal pyrolysis [J]. Journal of materials research and technology, 2014, 3(1): 2-8.
- [24] 刘宇洋,高文韩,孔令芝,等.化学添加剂对有氧胁迫下湿贮玉米发酵品质和有氧稳定性的影响[J].草地学报,2022,30(12):3455-3463.LIU Y Y, GAO W H, KONG L Z, et al. Effects of chemical additives on the fermentation quality and aerobic stability of high moisture corn under the air stress [J]. Acta agrestia sinica, 2022, 30 (12) : 3455-3463 (in Chinese with English abstract).
- [25] 沈孟飞,宋虎潮,邢定一,等.废旧轮胎热解吸附强化重整制氢特性及经济性分析[J].西安交通大学学报,2024,58(10): 222-232.SHEN M F, SONG H C, XING D Y, et al. Characteristics and economic analysis of hydrogen production by pyrolysis adsorption enhanced reforming of waste tires [J]. Journal of Xi'an JiaoTong University, 2024, 58(10): 222-232 (in Chinese with English abstract).

- [26] 陈泳兴.*Undaria pinnatifida* 残渣水热法资源化利用及产物中羟基羧酸分离技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2017. CHEN Y X. Study on hydrothermal resource utilization of *Undaria pinnatifida* residue and separation technology of hydroxycarboxylic acid in the product [D]. Harbin: Harbin Institute of Technology, 2017(in Chinese with English abstract).
- [27] DAI Y, LI C H, LIU J Q, et al. Enhancing the stability of mung bean-based milk: insights from protein characteristics and raw material selection[J/OL]. International journal of biological macromolecules, 2024, 265: 131030 [2024-07-04]. <https://doi.org/10.1016/j.ijbiomac.2024.131030>.
- [28] 武利斌, 常岩超, 唐宏彬, 等. 350MW 超临界对冲燃烧锅炉煤与生物质混燃数值模拟研究[J]. 青海电力, 2024, 43(2): 58-64. WU L B, CHANG Y C, TANG H B, et al. Numerical simulation of mixed combustion of coal and biomass in 350 MW supercritical offset combustion boiler [J]. Qinghai electric power, 2024, 43(2): 58-64(in Chinese with English abstract).
- [29] 桂标. 物化预处理对市政污泥水热碳化产物生成规律及重金属污染风险的影响[D]. 南京: 南京信息工程大学, 2023. GUI B. Effect of physicochemical pretreatment on the formation law of hydrothermal carbonization products of municipal sludge and the risk of heavy metal pollution [D]. Nanjing: Nanjing University of Information Science & Technology, 2023(in Chinese with English abstract).

Intelligent modeling and response behavior of process of producing hydrogen-rich hydrothermal bio-oil from wastes with high-moisture and high-lipid

LIU Xiangjie, YUAN Qiaoxia, CAO Hongliang

College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

Abstract Two types of high-fitness machine learning algorithms including random forest and extreme gradient enhancement Tree were used to establish a prediction model of the content of hydrogen in hydrothermal bio-oil with high-precision and wide-range ($R^2 > 0.93$) based on 243 sets of experimental data collected from literature on the hydrothermal preparation of hydrogen-rich bio-oil from wastes with high-moisture and high-lipid including dead livestock and algae to construct a model for predicting the content of hydrogen in hydrothermal bio-oil with high accuracy of prediction and strong ability of generalization, and to study in depth the laws and mechanisms of hydrothermal conversion of biomass. SHAP interpretable technology and local dependence analysis method were used to analyze the contribution, local dependence response behavior and interactive coupling law of conditions for producing hydrothermal bio-oil based on the mathematical prediction model. The results showed that the content of lipid and hydrogen in the wastes with high-moisture and high-lipid were the determining factors for the preparation of hydrogen-rich bio-oil, ranked the top two in terms of its contribution to the enrichment of hydrogen in the oil phase and significantly affected the accumulation of hydrogen in the oil phase. The content of hydrogen in the oil phase was increased with the increase of the content of hydrogen in the raw material, indicating that hydrogen-rich raw materials provide convenient conditions for the preparation of hydrogen-rich bio-oil, with an improvement effect of up to 4%. The high calorific value of raw materials was the main inhibitory factor for the hydrogen-enriched behavior of bio-oil, with an inhibitory effect of up to 4%. The coupling effect between the information of element, industry, and components of biomass was strong from the interactions among the characterized categories, but the effect of local coupling between the characteristics of raw material and operating conditions was relatively small.

Keywords data-driven; machine learning; random forest; extreme gradient boosting tree; bio-oil; calorific value

(责任编辑:陆文昌)