

基于机器学习的清洁氢度氢成本评估研究

郭文杰 杨景宇 杨元元

中国矿业大学(北京), 北京 100000

摘要: 在“双碳”目标背景下, 清洁氢作为零碳、高效、可大规模存储的二次能源, 是能源转型与工业脱碳关键方向。PEM电解槽因响应速度快、适配波动性可再生能源的特点, 成为清洁氢制备的主流技术路线。当前清洁氢度氢成本偏高、影响因素复杂、区域差异显著, 制约了产业规模化落地。为识别成本驱动机制, 本文以PEM电解槽制氢为研究对象, 选取内蒙古、青海、江苏、广东、上海5个典型地区2020—2024年面板数据, 构建轻量级机器学习成本评估模型, 以工业电价、电解槽单位投资、弃风弃光率、地方制氢补贴为输入特征, 以单位度氢成本(LCOH)为预测目标, 建立多元线性回归基准模型与随机森林增强模型, 通过模型对比、特征重要性分析与区域场景化预测, 解析清洁氢成本的影响规律与经济性潜力。结果表明, 多元线性回归模型 $R^2=0.99$, 精度良好; 随机森林模型特征重要性排序显示, 工业电价和电解槽投资为核心影响因素; 2025年内蒙场景预测成本为23.39元/kg, 与实际公开项目成本趋势一致。研究可为清洁氢降本路线、政策制定、区域布局提供支撑与参考。

关键词: 清洁氢; 度氢成本; 机器学习; 随机森林; 场景预测

DOI: 10.64649/yh.shygl.issn3105-0085.202605014

0 引言

在“双碳”目标背景下, 氢能凭借零碳高效成为我国能源转型与工业脱碳的关键载体^[1]。清洁氢由可再生能源电解水制备, 全生命周期零碳, 在多领域具有突出的应用价值^[2], 是未来能源体系的重要组成部分。PEM电解槽响应快、适配波动性电源, 是清洁氢规模化制备的重要技术路线^[3]。

当前我国清洁氢规模化发展的主要瓶颈源于度氢成本偏高, 经济效益不足^[4]。合理评估成本, 厘清关键影响因素, 对技术发展、政策制定与行业布局具有重要意义^[5]。本文采用轻量级机器学习结合传统成本测算思路, 可为地方差异化产业政策制定、企业投资选址规划提供支撑^[6]。

现有PEM制氢成本研究多采用LCOH模型静态测算, 依赖参数假设, 虽然方法成熟但难准确分析多因素的复杂影响^[4]。王彦哲等基于学习曲线的研究重点分析长期发展趋势^[5], 缺少地方的精细化测算。机器学习已用于能源成本预测, 适配小样本与多维度场景, 但现有研究多聚焦碱性电解, 针对PEM且面向多省份轻量化成本评估相对较少。

本文围绕PEM电解槽制氢成本展开研究, 依托国家能源统计数据、可再生能源发展报告和氢能产业实际资料, 梳理成本理论体系与关键影响因素, 选取多维度合理数据, 构建多元线性回归与随机森林双模型, 通过特征重要性排序解析成本驱动机制, 最后进行不同区域场景下的成本预测, 凝练结论建议。

1 理论基础与模型构建

1.1 清洁氢制氢成本理论基础

PEM电解槽以质子交换膜电解水制氢, 具备运行灵活、氢气纯度高、适配新能源发电的优势, 是清洁氢产业规模化发展的主流技术路线^[3]。

单位度氢成本是衡量清洁氢经济效益的核心指标, 具体指制氢项目在全生命周期内生产单位氢气的平均成本, 含电力、设备折旧、运维、补贴等, 电力占比超60%, 设备投资约20%~30%, 是两大关键项。

1.2 特征变量选取与数据源

本文选取内蒙古、青海、江苏、广东、上海5省2020—2024年数据, 共25组样本。研究区域覆盖西北新能源富集区与东部高负荷地区, 差异显著, 能较好反映我国清洁氢成本的区域特征。研究选取4类输入特征与1项目标变量。工业电价决定制氢可变成本; 电解槽单位投资影响成本长期分摊; 弃风弃光率体现新能源消纳压力, 高弃电地区可获取低成本电力; 地方制氢补贴可有效对冲成本压力。以清洁氢度氢成本为目标变量, 作为项目经济性的核心评估标准。数据来自中国氢能联盟白皮书等权威公开渠道^{[6][7]}, 多源数据交叉验证, 真实合理, 为研究分析提供基础。

1.3 机器学习评估模型设计

1.3.1 基础模型: 多元线性回归

多元线性回归模型简单、解释性强、运算快, 可量化边际影响, 适合作为基准模型拟合线性关系^[5]。数学表达式如下
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \#$$

其中 Y 为度氢成本 (元/kg); X_1-X_4 依次为工业电价 (元/kWh)、电解槽投资 (元/kW)、弃风弃光率 (%)、制氢补贴 (元/kg); β_0 为截距项, $\beta_1-\beta_4$ 为回归系数; ε 为服从正态分布的随机误差项。

1.3.2 增强模型: 随机森林

随机森林是集成模型, 具备较强的抗干扰能力与非线性拟合能力, 可输出特征重要性, 解析成本驱动机制。作为增强模型, 与多元线性回归形成对比。

基于 Python 的随机森林回归模型构建预测模型, 采用 100 棵决策树保证预测稳定, 设置随机数种子为 42 确保结果可复现, 以均方误差为损失函数、平方根为节点最大特征选取规则, 并启用自助采样提升泛化能力, 其余参数采用默认设置。

1.3.3 模型评价指标

本文采用以下指标衡量模型性能。平均绝对误差 (MAE), 反映预测值与真实值的平均绝对偏差, 数值越小, 预测精度越高; 拟合优度 (R^2), 反映模型对数据变异的解释能力, 取值范围 [0,1], 数值越接近 1, 拟合效果越好; 均方根误差 (RMSE), 可突出极端偏差, 数值越小, 模型稳定性越强。

2 结果与分析

2.1 数据预处理与描述性统计

数据经清洗、去异常与标准化处理, 保证输入质量。样本共 25 组, 涵盖 5 个省份 2020-2024 年面板数据。

表1 变量描述性统计

变量	样本量	均值	标准差	最小值	中位数	最大值
工业电价	25	0.52	0.16	0.30	0.60	0.70
电解槽投资	25	4160	810.35	3000	4200	5200
弃风弃光率	25	5.42	5.76	0.50	1.50	16.10
制氢补贴	25	0.52	0.16	0.20	0.50	0.80
度氢成本	25	30.70	5.87	20.20	30.40	41.30

为验证变量间关系与多重共线性, 绘制相关性热力图。图 1 结果显示, 度氢成本与工业电价与电解槽投资呈显著正相关; 与弃风弃光率呈显著负相关, 符合成本经济学规律。工业

电价与弃风弃光率强负相关, 其余自变量间相关系数均小于 0.5, 无严重多重共线性, 可用于模型构建。

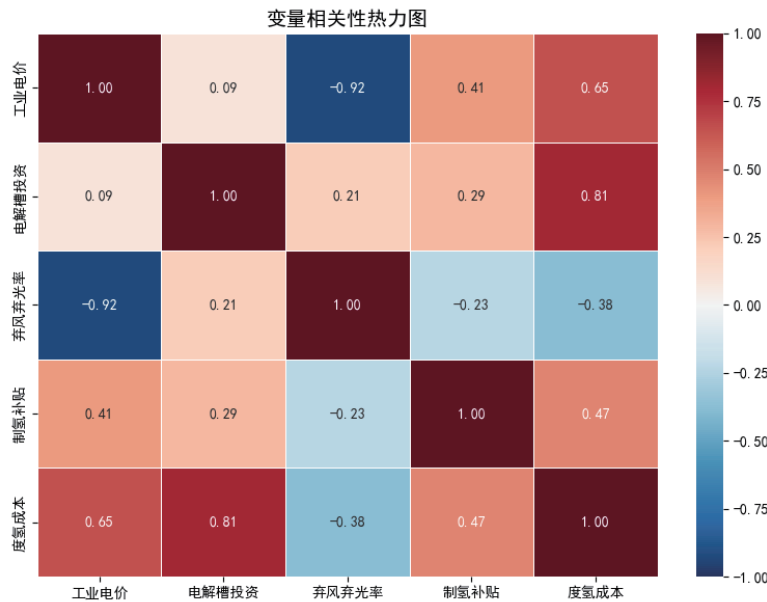


图1 变量相关性热力图

2.2 模型训练与拟合效果验证

按 8:2 划分训练集与测试集构建模型, 精度对比见表 2。

表2 预测精度对比

模型	MAE (元/kg)	R^2	RMSE (元/kg)
多元线性回归	0.32	0.99	0.41
随机森林	1.28	0.88	1.56

由表3-2可知,多元线性回归预测精度高, R^2 接近1,MAE与RMSE极低,说明核心影响因素与度氢成本呈显著线性关系。随机森林精度略低,但指标合理,非线性拟合能力较强,适合计算特征重要性。两模型结合使用可实现预测与影响机制解析,为场景化分析提供支撑。

2.3 特征重要性与驱动机制分析

随机森林特征重要性依次为:工业电价51.4%,电解槽投资36.0%,制氢补贴7.1%,弃风弃光率5.5%。电价与设备投资主导项目经济性,补贴与弃风弃光率有辅助调节作用。结论与产业一致,可为降本路径提供方向。

2.4 区域清洁氢成本场景化预测

本文选取内蒙古、青海、江苏开展2025年场景化预测。结合技术降本、新能源平价与政策导向,设定电价下降等合理假设。

结果显示:2025年内蒙清洁氢度氢成本预测值为23.39元/kg,青海为22.87元/kg,江苏为31.52元/kg。预测值略高,偏差源于模型采用区域平均工业电价,而实际项目可获长期协议电价,高设备利用率与专项补贴。但成本量级与变化趋势与实际一致,西部25元/kg以下,东部30元/kg以上,预测结果可靠。

3 结论与展望

3.1 研究主要结论

本文构建双模型,实现PEM电解槽制氢成本精准预测。其中多元线性回归模型的预测精度稳定, $R^2=0.99$,而随机森林具备小样本、低维度、高可解释性特点。

特征重要性分析表明,工业电价与电解槽

投资是影响PEM电解槽制氢成本的核心因素,二者合计贡献度达87.4%,共同决定项目经济性,补贴与弃风弃光率起辅助调节作用。

区域对比显示,各省度氢成本差异明显,原因在于资源禀赋、电价与设备投资水平。2025年场景预测表明,内蒙古、青海等西部新能源富集区成本可进一步下降,趋势与实际项目基本一致。

3.2 政策建议与产业启示

核心成本方面。可依托绿电直购、长期协议电价等政策降低用电成本;同时扶持PEM电解槽研发、核心材料国产化与规模化生产,削减设备投资。

针对区域差异问题。西北富集区完善弃风弃光消纳机制,落实电价优惠,提升电解槽利用率;东部地区加大设备与应用补贴,兼顾技术降本与就近消纳。

针对补贴,实现精准倾斜。将补贴重点向资源优势区、消纳困难区倾斜,提升政策效能,避免补贴资源分散。

最重要的是建立成本动态监测体系。用轻量机器学习模型搭建成本监测平台,定期跟踪成本与因素变化,用监测报告支撑政策、布局与投资决策。

3.3 研究不足与展望

本文存在一定局限:样本范围有限,没有考虑电解槽效率、运维成本的影响,模型以轻量为主。未来可扩大样本覆盖,补充因素,完善敏感性分析和多技术路线对比,由此研究的广度、深度与实用价值得到提升。

参考文献:

- [1] 程文姬,赵磊,郝航,等.“十四五”规划下氢能政策与电解水制氢研究[J].热力发电,2022,51(11):181-188.
- [2] 何泽兴,史成香,陈志超等.质子交换膜电解水制氢技术的发展现状及展望[J].化工进展,2021,40(09):4762-4773.
- [3] 林旗力,戚宏勋,黄晶晶,等.碱性-质子交换膜水电解复合制氢平准化成本分析[J].储能科学与技术,2023,12(11):3572-3580.
- [4] 刘玮,万燕鸣,熊亚林.碳中和目标下电解水制氢关键技术及价格平准化分析[J].电工技术学报,2022,37(11):2888-2896.
- [5] 王彦哲,欧训民,周胜.基于学习曲线的中国未来制氢成本趋势研究[J].气候变化研究进展,2022,18(03):283-293.
- [6] 国家能源局.中国可再生能源发展报告[R].北京:国家能源局,2024.
- [7] 中华人民共和国国家统计局.中国统计年鉴2024[M].北京:中国统计出版社,2024.

作者简介:郭文杰(2005.06—),男,汉族,北京市人,本科在读,研究方向为大数据管理与应用、能源经济与机器学习。

项目信息:中央高校基本科研业务费资助和中国矿业大学(北京)大学生创新训练项目资助(202505014)。